

Dissertation zur Erlangung des Doktorgrades
der Fakultät für Chemie und Pharmazie
der Ludwig-Maximilians-Universität München

Synthese chemisch modifizierter, epigenetisch relevanter Nukleoside und Oligonukleotide

Alexander Maximilian Schön

aus

München, Deutschland

2020

Erklärung

Diese Dissertation wurde im Sinne von § 7 der Promotionsordnung vom 28. November 2011 von Herrn Prof. Dr. Thomas Carell betreut.

Eidesstaatliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, den 07. Oktober 2020

Alexander Schön

Dissertation eingereicht am: 09. Oktober 2020

1. Gutachter: Prof. Dr. Thomas Carell

2. Gutachterin: Dr. Sabine Schneider

Mündliche Prüfung am: 23. November 2020

„Life is the art of drawing without an eraser“

John W. Gardener

*Für meine Eveliina
und unseren kommenden Nachwuchs*

Teile dieser Arbeit wurden bereits in Fachjournalen publiziert oder auf Fachkonferenzen vorgestellt.

Publikationsliste

A. Schön[#], E. Kaminska[#], F. Schelter[#], E. Ponkkonen, E. Korytiaková, Sarah Schiffers, T. Carell^{*}; Analysis of an active deformylation mechanism of 5-formyl-deoxycytidine (fdC) in stem cells, *Angew. Chem. Int. Ed.* **2020**, 59(14), 5591-5594; DOI: 10.1002/anie.202000414.

R. C. A. Dubini, A. Schön, M. Müller, T. Carell, P. Rovó^{*}; Impact of 5-formylcytosine on the melting kinetics of DNA by ¹H NMR chemical exchange, *Nucleic Acids Res.* **2020**, 48(15), 8796-8807, DOI: 10.1093/nar/gkaa589.

P. Thumbs[#], T. T. Ensfelder[#], M. Hillmeier[#], M. Wagner[#], M. Heiss, C. Scheel, A. Schön, M. Müller, S. Michalakakis, S. Kellner, T. Carell^{*}; *Angew. Chem. Int. Ed.* **2020**, 59(30), 12352-12356, DOI: 10.1002/anie.202002295.

Konferenzbeiträge

SFB 1309 Meeting, München 2017 – Posterpräsentation „*Novel Chemically-Enabled Sequencing Methods for DNA Modifications*“

Inhaltsverzeichnis

1. Zusammenfassung	I
2. Summary	III
3. Einleitung.....	1
3.1. Epigenetik und DNA-Methylierung.....	1
3.2. Inhibition der Methylierung durch 5-Azacytidin	2
3.3. Oxidierte mdC-Derivate und Mechanismen der Demethylierung.....	5
3.4. Methoden zur Sequenzierung der epigenetisch relevanten Cytosin- Modifikationen	9
3.5. Beispiele von RNA-Modifikationen: Inosin und Queuosin.....	13
4. Aufgabenstellung.....	16
5. Veröffentlichte Arbeiten	20
5.1. Analysis of an active deformylation mechanism of 5-formyl-deoxycytidine (fdC) in stem cells.....	20
5.2. 5-Formylcytosine enhances dsDNA kinetics by globally destabilizing the double stranded structure	65
5.3. Synthesis of galactosyl-queuosine and distribution of hypermodified Q- nucleosides in mouse tissues	79
6. Unveröffentlichte Arbeiten	86
6.1. Weitere 6-Aza-substituierte Analoga der epigenetisch relevanten dC- Modifikationen	86
6.1.1. Synthesen und Eigenschaften von 6-Aza-mdC und 6-Aza-hmdC.....	86
6.1.2. Projektbeiträge	92
6.1.3. Experimentalteil	93
6.2. Untersuchung der direkten Demethylierung mittels eines isoto- penmarkierten dC-Derivats.....	102
6.2.1. Synthese der Verbindung und erste Ergebnisse.....	102
6.2.2. Chemische Derivatisierung zur Verbesserung der Signalintensität.....	111
6.2.3. Projektbeiträge	115
6.2.4. Experimentalteil	116
6.3. Untersuchungen der Aktivität von Sirtuinen auf die Deformylierung von fdC.....	129
6.3.1. Synthese eines DNA-Oligonukleotids mit isoto- penmarkiertem fdC und <i>in vitro</i> -Experimente mit hSirt1	129
6.3.2. Projektbeiträge	134

6.3.3.	Experimentalteil	135
6.4.	Entwicklung von Hydroxylamin-Reagenzien zur Sequenzierung von genomischem fdC mittels Technologien der 3. Generation	143
6.4.1.	Design und Synthese der Verbindungen	143
6.4.2.	Reaktion der Derivatisierungsreagenzien mit fdC-enthaltenden Oligonukleotiden	153
6.4.3.	Zusammenfassung und Ausblick	154
6.4.4.	Projektbeiträge.....	155
6.4.5.	Experimentalteil	156
6.5.	Synthese eines carbazyklischen Derivats von 5-Azacytidin und Untersuchungen zu dessen Einfluss auf <i>in vivo</i> Methylierung von DNA....	176
6.5.1.	Synthese eines carbazyklischen 5-Azacytidins.....	176
6.5.2.	Untersuchungen zur Hydrolysestabilität der carbazyklischen Verbindung.....	179
6.5.3.	Zellverfütterungsexperimente mit c-AzaC und erhaltene Ergebnisse .	181
6.5.4.	Synthese eines Nukleotid- <i>Prodrugs</i> von cAzaC und Untersuchung dessen in Zellverfütterungsexperimenten	183
6.5.5.	Projektbeiträge.....	187
6.5.6.	Experimentalteil	188
6.6.	Untersuchung der Basenpaarungseigenschaften von Inosin mittels eines isotoopenmarkierten Inosin-RNA-Duplexes	201
6.6.1.	Synthese eines ¹⁵ N-isotoopenmarkierten Inosin-Phosphoramidits, Oligonukleotidsynthese und erste NMR-spektroskopische Ergebnisse	201
6.6.2.	Projektbeiträge.....	206
6.6.3.	Experimentalteil	207
7.	Material und Methoden.....	214
8.	Abkürzungsverzeichnis.....	218
9.	Literaturverzeichnis	223
10.	Danksagung	228

1. Zusammenfassung

Epigenetische Modifikationen von Nukleosiden haben großen strukturellen sowie funktionellen Einfluss sowohl auf DNA als auch auf RNA in lebenden Organismen. Während in der RNA viele verschiedene und auch in ihrer Komplexität stark variierende Modifikationen vorkommen können, beschränken sich DNA-Modifikationen großteils auf die 5-Position von Cytosin-Einheiten. Neben der gut verstandenen Methylierung dieser Position unter Ausbildung von 5-Methyl-2'-desoxycytidin (mdC) sind auch weitere enzymkatalysierte Oxidationsreaktionen von mdC zum 5-Hydroxymethyl-2'-desoxycytidin (hmdC), 5-Formyl-2'-desoxycytidin (fdC) und 5-Carboxyl-2'-desoxycytidin (cadC) bekannt.

Insbesondere der Einfluss der bisher nicht ausreichend untersuchten fdC-Modifikation auf die DNA-Doppelstrangstruktur sowie der bisher unbekannte Mechanismus deren aktiver Entfernung aus genomischer DNA wurden im Rahmen dieser Arbeit intensiv untersucht. Dazu wurden chemisch modifizierte Nukleoside und Oligonukleotide synthetisiert, die in Experimenten sowohl *in vivo* als auch *in vitro* unter Verwendung moderner Analysetechniken eingesetzt und untersucht wurden. Dadurch konnten eindeutige mechanistische sowie quantitative Rückschlüsse zur Aufklärung der aktiven Demethylierung von mdC über die Oxidation zu fdC in genomischer DNA erhalten sowie der strukturelle Einfluss dieser Modifikation auf doppelsträngige DNA weiter aufgeklärt werden. Da auch die sequenzspezifische Detektion dieser Modifikation entscheidend zum Verständnis ihrer Funktion beiträgt, wurden im Rahmen dieser Arbeit spezifische Derivatisierungsreagenzien entwickelt, die zur Weiterentwicklung von Sequenzierungstechniken zur Detektion von genomischem fdC verwendet werden könnten.

Des Weiteren wurde ein hydrolysestabiles Derivat des bekannten, Methylierungsinhibitors Azacitidin entwickelt, welches in Zellkulturen eingesetzt in seiner Form als Nukleotid-Monophosphat-Vorläufer einen hypomethylierenden Effekt auf die DNA der Zellen bei gleichzeitig erhöhter Hydrolysestabilität aufwies. Demzufolge könnte dies einen interessanten Ansatzpunkt zum Ersatz der wasserempfindlichen Verbindung in der klinischen Anwendung darstellen.

Schließlich wurden auch die RNA-Modifikationen Inosin und Galaktosylqueuosin untersucht, wobei zum einen das Basenpaarungsverhalten der Inosinbase mit Uridin

an Hand von isotoopenmarkierten, doppelsträngigen RNA-Oligonukleotiden NMR-spektroskopisch analysiert wurde. Zum anderen konnte die exakte regio- sowie stereochemische Struktur der Modifikation Galaktosylqueuosin anhand von Koinjektionsexperimenten eines synthetisch hergestellten Standards mit dem natürlichen Nukleosidgemisch aufgeklärt werden.

2. Summary

Epigenetic modification of nucleosides has strong structural as well as functional impact on DNA and RNA of living organisms. Whereas a big variety of modifications differing in their structural complexity may occur in RNA, DNA modifications are mainly occurring at the 5-position of genomic cytosine moieties. Further than the well understood methylation of said position to form 5-methyl-2'-deoxycytidine (mdC) enzymatic oxidation reactions are known to further oxidize mdC to 5-hydroxymethyl-2'-deoxycytidine (hmdC), 5-formyl-2'-deoxycytidine (fdC) and 5-carboxy-2'-deoxycytidine (cadC).

Particularly, the structural impact of the not yet fully understood fdC modification on DNA double-strand structures and also the mechanism regarding its removal via the so-called pathway of direct, active demethylation were investigated thoroughly in the course of this work. To this aim, chemically modified nucleosides and oligonucleotides were synthesized and used in studies applying modern analytical detection techniques, both *in vivo* or *in vitro*. As a result, clear mechanistic and quantitative insight was gained to understand the process of active demethylation of mdC via oxidation to fdC in genomic DNA and, furthermore, the structural impact of said modification on DNA double-strand structures was clarified applying an artificial oligonucleotide model system. As, moreover, sequence specific detection of the fdC modification is of interest to further understand this modification, new derivatization reagents were synthesized in the course of this work which could lead to further improvements in sequencing protocols for genomic fdC.

Further, a hydrolytically stable derivative of the known water labile methylation inhibitor azacitidin was synthesized, which indeed resulted upon administration as nucleotide monophosphate precursor in a hypomethylating effect on the DNA in cell culture simultaneously maintaining its stability towards hydrolysis. This could be a starting point for substitution of the known water labile compounds in clinical application.

Furthermore, the RNA modifications inosine and galactosyl queuosine were investigated, whereby the base pairing properties of inosine nucleosides paired with uridines in double-stranded, isotopically labeled RNA oligonucleotides were investigated via NMR spectroscopy. Moreover, the exact regiochemical and

stereochemical properties of galactosyl queuosine were elucidated by comparing a synthetic standard with natural nucleoside mixtures in coinjection experiments.

3. Einleitung

3.1. Epigenetik und DNA-Methylierung

Ausdifferenzierte Zellen eines multizellulären Organismus, die alle von einer multipotenten Stammzelle abstammen, können sich stark in ihrem jeweiligen Proteom, d.h. in der Gesamtheit der in der jeweiligen Zelle vorkommenden Proteine, unterscheiden.^[1] Um dies ausgehend von der identischen genetischen Information der Zellen umsetzen zu können, müssen diese Zellen die Fähigkeit besitzen, gezielt bestimmte Gene zu definierten Zeitpunkten aktivieren oder deaktivieren zu können.^[1,2] Dies ist nicht nur für die Ausbildung des gewebetypischen Proteoms der ausdifferenzierten Zellen sondern darüber hinaus auch für die Reaktion der Zellen auf exogene Einflüsse unabdingbar.^[1,2] Diese Beeinflussung der Genexpression zur Ausbildung eines Zelltyp-spezifischen Proteoms wird der Epigenetik zugeordnet.^[1] Die epigenetische Genregulation wird unter anderem durch kovalente DNA-Modifikationen sowie durch Modifikationen an den Histonen, um die die DNA zur Größenkomprimierung gewickelt ist, umgesetzt.^[3,4]

Die häufigste kovalente DNA-Modifikation ist die Methylierung der C5-Position von 2'-Desoxycytidin (dC) zum 5-Methyl-2'-desoxycytidin (mdC)^[2,5], die innerhalb des menschlichen Genoms an bis zu 4 % und in murinen embryonalen Stammzellen (mESCs) an bis zu 6 % der vorkommenden Cytosin-Basen auftreten kann.^[2,5] Diese Modifikation ist mehrheitlich auf CpG-Dinukleotiden benachbart zu einem Guanin lokalisiert.^[6–8] Aufgrund der Symmetrie eines solchen CpG-Dinukleotids innerhalb des DNA-Doppelstranges ist eine Methylierung beider DNA-Stränge und somit eine Beibehaltung der Modifikation bei der Replikation möglich.^[2] Aus diesem Grund kann mdC auch als epigenetische Langzeitmodifikation klassifiziert werden.^[9] Die Präsenz von genomischem mdC kann verschiedene Auswirkungen haben. So kann mdC neben der Repression der Genexpression, bei der es das Binden von Transkriptionsfaktoren moduliert^[10,11] oder spezielle mdC-bindende Proteine bindet, die dann ihrerseits Korepressor-Komplexe anreichern,^[3] auch weitere Möglichkeiten der Einwirkung auf die Genexpression haben. Bekannt sind unter anderem Einflüsse auf X-Chromosom-Desaktivierung,^[12] Genom-*Imprinting*^[13,14] und Stilllegung von Transposonen.^[15,16] Wegen der vielfältigen essentiellen Effekte der DNA-Methylierung auf die zellulären Prozesse, kann eine fehlerhafte oder gar fehlende DNA-Methylierung mit Krebserkrankungen, psychischen Störungen oder

Erkrankungen, die mit genomischer Prägung zusammenhängen, in Verbindung gebracht werden.^[17–19]

Auf chemischer Ebene betrachtet wird die enzymatische Methylierung von genomischem dC durch S-Adenosylmethionin-(SAM)-abhängige Methyltransferasen (*DNA nucleotide methyltransferases*, Dnmts) katalysiert, wobei innerhalb der Dnmts verschiedene Vertreter der Methyltransferasen vorkommen. Dnmt1 ist eine sog. Erhaltung-Methyltransferase (*maintenance methyltransferase*), die nach der Replikation mit hoher Spezifität die nun hemimethylierten DNA-Stränge komplementär methyliert, um mdC als epigenetische Information zu erhalten.^[20,21] Im Gegensatz dazu zeigen Dnmt3a und Dnmt3b keine besondere Spezifität für hemimethylierte oder unmethylierte DNA, weswegen diese als *de novo* Methyltransferasen bezeichnet werden.^[21,22] Unabhängig vom Typ der Dnmt verläuft der bereits gut verstandene Mechanismus der Methylierung vom dC zum mdC grundsätzlich nach einem generellen mechanistischen Muster (siehe Abbildung 3-1).^[15]

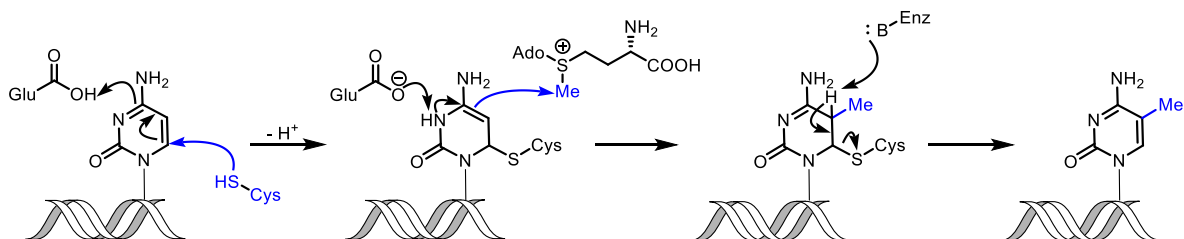


Abbildung 3-1: Mechanismus der Methylierung von dC zum mdC im Genom durch ein Dnmt-Enzym.

Der initiale Schritt der Methylierung ist dabei ein nukleophiler Angriff einer enzymatischen Cystein-Seitenkette der Methyltransferase an die 6-Position des Pyrimidinrings von dC. Anschließend erfolgt die Methylierung der 5-Position, wobei SAM die Methylgruppe unter Freisetzung einer Adenosylhomocysteineinheit überträgt. Eine β -Eliminierung des Thio-Nukleophils nach Deprotonierung an der 5-Position unter Rearomatisierung der Nukleobase und gleichzeitiger Regeneration des Enzyms liefert schließlich mdC.

3.2. Inhibition der Methylierung durch 5-Azacytidin

Die Kontrolle der DNA-Methylierung mit Hilfe von Nukleosid-Analoga stellt insbesondere im Hinblick auf deren Verwendung als Chemotherapeutika einen interessanten Ansatzpunkt für die medizinisch-chemische Forschung dar. So ist bekannt, dass die Verabreichung von 5-Aza-2'-desoxycytidin (Decitabin, Aza-dC)

oder dessen Ribose-Analogon 5-Azacytidin (Azacitidin, Aza-C) einen hypomethylierenden Effekt auf die DNA des behandelten Organismus hat.^[23–26] Diese Kontrolle der epigenetischen Information durch Decitabin und Azacitidin wird bereits zur Behandlung von akuter myeloischer Leukämie (AML)^[27] oder von myelodysplastischen Syndromen (MDS)^[28–30] klinisch eingesetzt. Dabei wird der hypomethylierende Effekt und möglicherweise die damit einhergehende Cytotoxizität der 5-Aza-Nukleoside ausgenutzt, um die Reaktivierung von durch mdC desaktivierten Tumorsuppressorgenen zu begünstigen.^[31]

Der chemische Wirkmechanismus der 5-Azacytidine beruht dabei auf der irreversiblen Inhibition der Dnmts.^[25,31] Nach dem nukleopilen Angriff der Cystein-Seitenkette im aktiven Zentrum des Enzyms an die 6-Position des Nukleosid-Derivats kann zwar eine Methylgruppe auf die 5-Position übertragen werden, eine β -Eliminierung des Enzyms ist im Falle der 5-Aza-Nukleoside allerdings nicht mehr möglich. Das so gebildete, kovalente Dnmt-DNA-Addukt kann somit nicht gespalten und die aktive Methyltransferase nicht rückgebildet werden (siehe Abbildung 3-2).^[32]

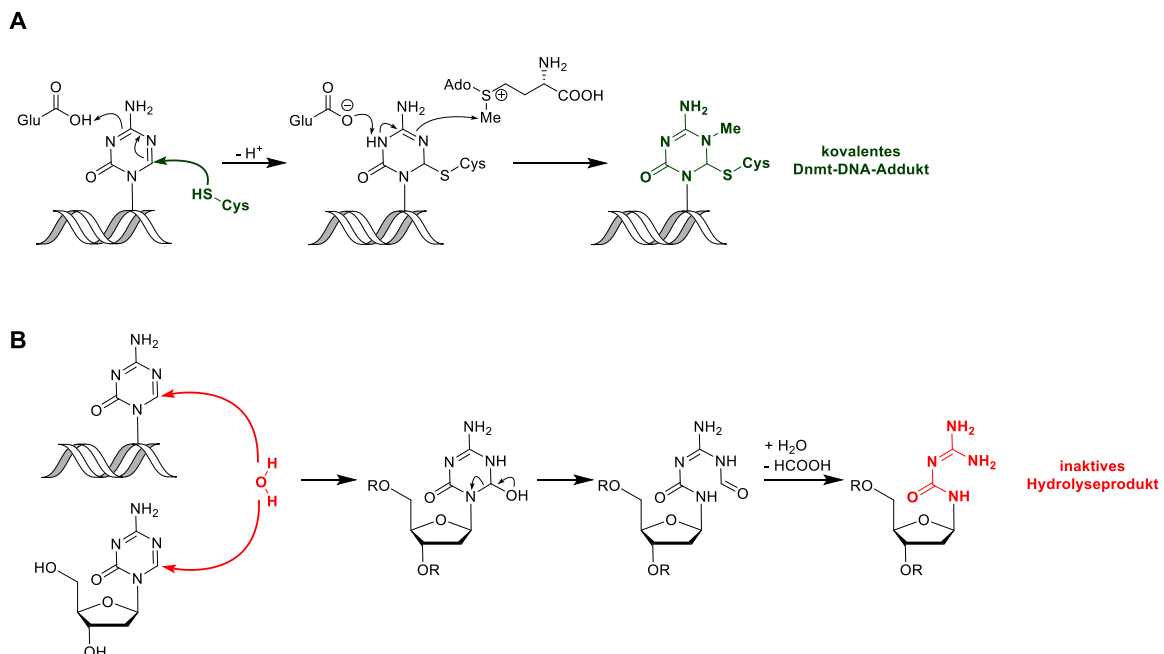


Abbildung 3-2: A) Wirkmechanismus der 5-Azacitidine, der zu einem kovalenten DNA-Dnmt-Addukt führt. B) Hydrolytische Zersetzung von freiem oder genomischem Aza-dC.

Auch für das Ribose-Analogon 5-Aza-C kann vermutet werden, dass dieses *in vivo* auf Nukleotidebene an der 2'-Position desoxygeniert und in der Folge in die DNA eingebaut werden kann, um dort seinen inhibierenden Effekt auf die Methyltransferasen wirken zu lassen.

Ein großer Nachteil bei der klinischen Verabreichung sowie der Handhabung der 5-Aza-Cytidine ist allerdings deren Anfälligkeit zur Hydrolyse in wässriger Umgebung. So kann ein konkurrierender O-Angriff eines Wassermoleküls an der 5-Position zur Hydrolyse der freien oder genomischen 5-Aza-Nukleoside führen.^[33] Dabei kommt es nach Wasser-Addition zur Ringöffnung des Triazinrings und anschließend zur irreversiblen Deformylierung. Die daraus resultierende offenkettige Guadinium-ähnliche Struktur zeigt in der Folge keinen Methyltransferase-inhibierenden Effekt mehr. Diese Hydrolyseanfälligkeit der 5-Azacytidine führt dazu, dass bei der klinischen Verabreichung der Medikamente eine höhere Dosierung erfolgen muss, um den Verlust durch die Hydrolyse abzufedern. Darüber hinaus kommt es auch zu einem ungleich höheren Verbrauch des eingesetzten Medikaments, da die zu verabreichende Lösung immer frisch angesetzt werden muss.^[34]

Erst kürzlich zeigten *Wildenhof et al.*^[32], dass durch den Austausch der 2'-Desoxyriboseeinheit hin zu einem Cyclopentan-basierten carbazyklischen Nukleosid die Hydrolyseneigung des Triazinrings deutlich vermindert werden konnte.^[32] Im Rahmen der Studien wurde die Stabilität des carbazyklischen 5-Aza-2'-desoxycytidins (c-Aza-dC) in wässrigem Kaliumhydrogenphosphatpuffer bei verschiedenen physiologischen pH-Werten untersucht und festgestellt, dass im Verlauf von zwei Wochen bei keinem der untersuchten pH-Werte eine hydrolytische Zersetzung der Verbindung zu beobachten war.^[32] Insbesondere die Langzeitstabilität des Moleküls gegenüber der Hydrolyse bei einem pH-Wert von 5.5 ist hierbei hervorzuheben, da bekannt ist, dass Tumorzellen oft einen leicht sauren pH-Wert aufweisen.^[35] Denn in derselben Studie konnte ebenfalls gezeigt werden, dass Aza-dC in saurer Umgebung bei einem pH-Wert von 5.5 besonders zur hydrolytischen Zersetzung neigt.^[32] Der hypomethylierende Effekt von c-Aza-dC blieb trotz der veränderten Hydrolyseeigenschaften des carbazyklischen Derivats bestehen.^[32] Anhand von Verfütterungsexperimenten von Aza-dC und c-Aza-dC an mESCs konnte gezeigt werden, dass allerdings eine höhere Dosis und längere Wirkdauer im Falle der Gabe von c-Aza-dC nötig ist, um einen mit Decitabin vergleichbaren Effekt zu erzielen.^[32] Die etwas geringere Wirksamkeit des carbazyklischen Derivats wird dabei auf dessen geringere Akzeptanz durch die zur Phosphorylierung notwendigen Kinasen zurückgeführt.^[32]

Alles in allem belegt die Studie, dass eine gezielte chemische Funktionalisierung, in diesem Fall die Verwendung eines Carbazyklus an Stelle einer 2'-Desoxyriboseeinheit, selektiv eine unerwünschte Eigenschaft einschränken, aber gleichzeitig einen gewünschten Effekt, wenn auch mit etwas geringerer Wirksamkeit, erreichen kann. Aza-dC und c-Aza-dC sind eindrucksvolle Beispiele, die zeigen, wie die epigenetische Information in Form von mdC anhand des Einsatzes von Nukleosid-Analoga *in vivo* gezielt beeinflusst werden kann.

3.3. Oxidierte mdC-Derivate und Mechanismen der Demethylierung

Neben der Methylierung von dC zu mdC wurde später auch entdeckt, dass mdC durch spezifische Fe(II)/ α -Ketoglutarat abhängige Dioxygenasen, die sog. *ten-eleven translocation* (TET)-Enzymfamilie,^[36–39] schrittweise über 5-Hydroxymethyl-2'-desoxycytidin^[36,40] (hmdC) zum 5-Formyl-2'-desoxycytidin^[5] (fdC) und schließlich zum 5-Carboxyl-2'-desoxycytidin^[37,41] (cadC) oxidiert werden kann. Die detektierbare Häufigkeit der oxidierten Derivate relativ zu unmodifiziertem dC nimmt dabei vom hmdC über das fdC bis hin zum cadC stetig ab.^[37,42] So lässt sich hmdC in verschiedenen Mausorganen mit 0.05% bis 0.15% aller dCs quantifizieren,^[43–45] wobei insbesondere im Gehirn ein deutlich höherer Gehalt an hmdC, bis hin zu 1.5% pro dC, detektiert werden kann.^[46,47] Für das aus hmdC resultierende fdC können Häufigkeiten von ca. 0.02% pro dC und für das noch höher oxidierte cadC von 0.002% pro dC ermittelt werden, die allerdings stark gewebeabhängigen Schwankungen unterliegen.^[5,48] Trotz der geringen Abundanz dieser Modifikationen konnten Verfütterungsexperimente mit isopenmarkiertem Material die Langlebigkeit von genomischem fdC *in vivo* demonstrieren, was zu der Schlussfolgerung führte, dass fdC eine funktionelle Rolle in der DNA haben könnte.^[49] Denn seit der Entdeckung der oxidierten dC-Derivate wird insbesondere deren Beteiligung an der aktiven Demethylierung von genomischem mdC, d.h. einer potentiellen Reaktivierung abgeschalteter Gene, untersucht. Die aktive Demethylierung wird als Gegenstück zur bekannten passiven Demethylierung gesehen, bei der mdC als epigenetische Information verloren geht, indem durch inaktive Methyltransferasen hemimethylierte DNA nach der Replikation nicht wieder komplementär methyliert wird (siehe Abbildung 3-3). Im Zuge mehrerer Zellteilungen führt dies schließlich zum Verlust der epigenetischen Information.

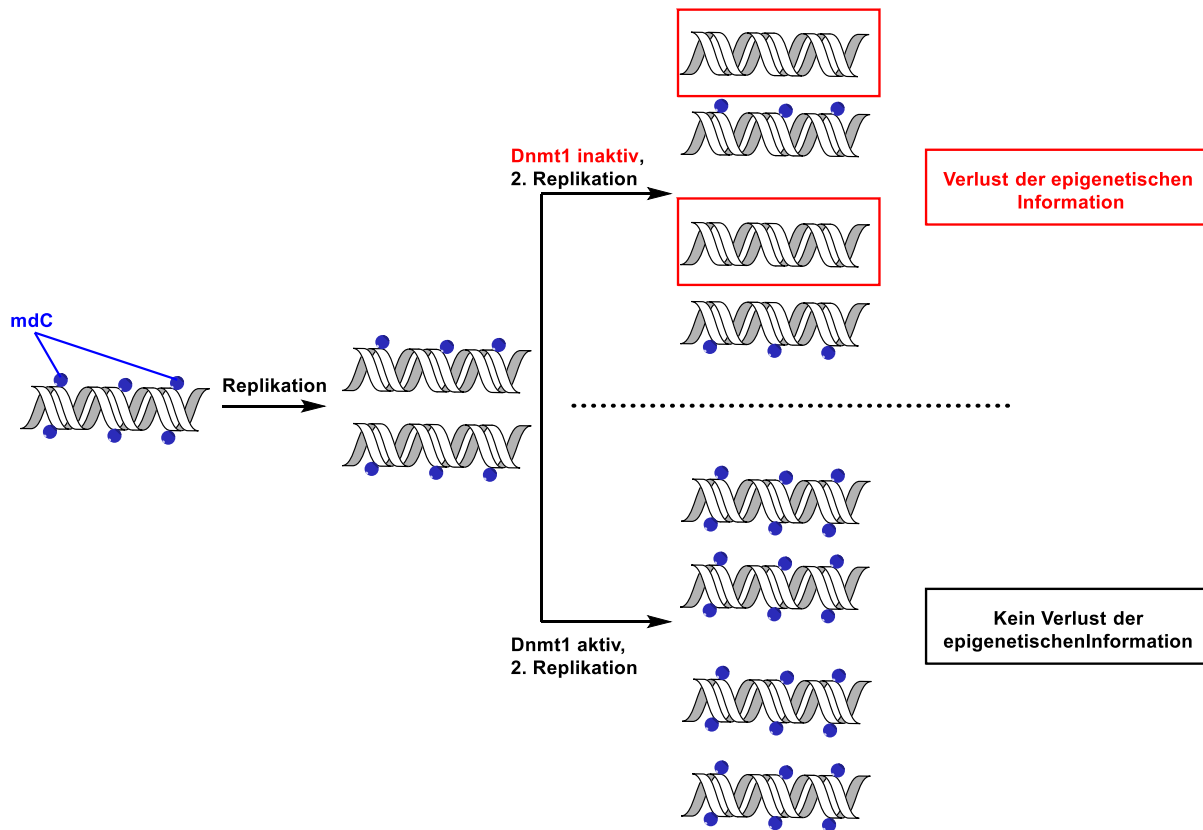


Abbildung 3-3: Vereinfachtes Schema der passiven Demethylierung, die im Verlauf mehrerer Replikationen zum Verlust der epigenetischen Information führt.

Die aktive Demethylierung hingegen wird als Prozess betrachtet, bei dem mdC direkt und selektiv aus dem Genom entfernt wird. Dabei werden die oxidierten Derivate fdC und cadC als Schlüsselmoleküle betrachtet.^[50] Zum einen ist bekannt, dass die sog. Thymin-DNA-Glycosylase (TDG) fdC und cadC als Substrate erkennen und durch Spaltung der glykosidischen Bindung die substituierte Nukleobase entfernen kann.^[41,51,52] Dieser Prozess der Basenexzisionsreparatur (BER) resultiert in einer sog. abasischen Stelle (*abasic site*) in der DNA. Anschließend wird die 2'-Desoxyriboseeinheit über eine β -Eliminierung des 3'-Phosphats unter Ausbildung eines Einzelstrangbruchs ausgeschnitten.^[53] Dieser Einzelstrangbruch kann im Anschluss durch den Einbau eines neuen Nukleotids und Ligation durch eine Polymerase repariert werden.^[54] Neben der sog. *short patch* Reparatur, bei der nur ein Nukleotid nach der Entstehung einer abasischen Stelle substituiert wird, kann auch eine *long patch* Reparatur erfolgen, bei der mehr als ein Nukleotid durch die DNA Polymerase ausgetauscht wird.^[54]

Neben der TDG-katalysierten BER, die fdC und cadC als Substrate erkennt, ist diese aber auch im Falle einer T-G-Fehlpaarung denkbar. Dabei kann mdC über eine

enzymatische Desaminierungsreaktion, die durch ein AID/APOBEC-Cytidineaminasesystem^[9,52] oder auch durch Dnmts^[55] katalysiert werden kann, in 2'-Desoxythymidin (dT) umgewandelt werden. Dabei entsteht eine fehlerhafte T-G-Basenpaarung, die durch Glykosylasen wie TDG oder MBD4 erkannt und mittels Exzision der fehlerhaften Base repariert werden kann.^[9] Darüber hinaus erkennt TDG auch 5-Hydroxymethyl-2'-desoxyuridin (hmdU) als Substrat für die BER.^[52,56] Eine Desaminierung von hmdC zu hmdU durch das oben genannte AID/APOBEC-Desaminasesystem kann durch spezielle hmdU-Glykosylasen wie SMUG1^[57] erkannt und mittels BER repariert werden.^[9] Die TET-katalysierte Oxidation von dT zu hmdU ist ebenfalls bekannt^[58] und bietet einen weiteren Weg der Demethylierung nach Desaminierung über hmdU (siehe Abbildung 3-4).

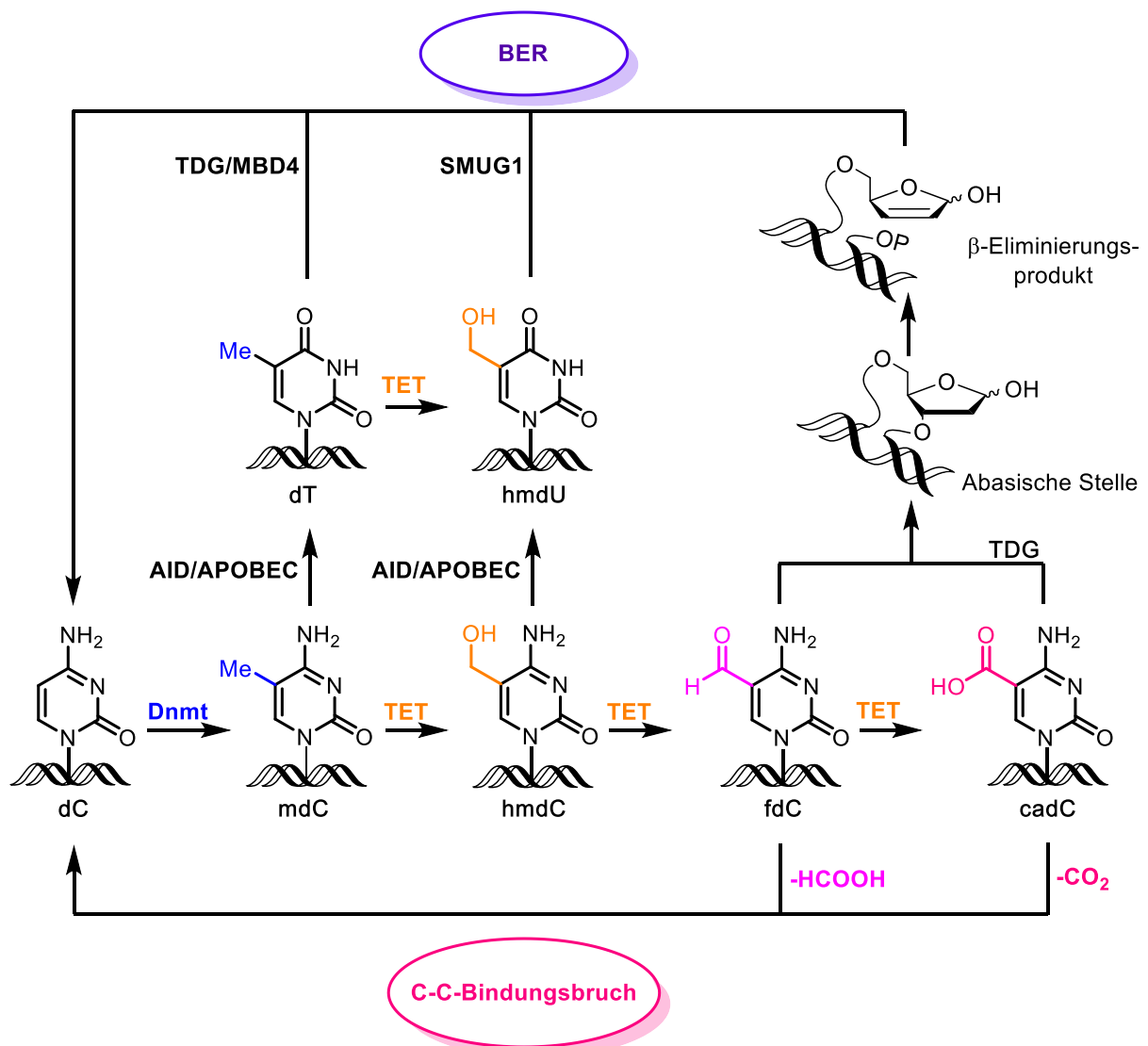


Abbildung 3-4: Darstellung der verschiedenen Wege der aktiven Demethylierung von mdC über BER (oben) oder direkte Demethylierung über C-C-Bindungsbruch (unten).

Neben der aktiven Demethylierung über BER konnte kürzlich gezeigt werden, dass auch ein TDG-unabhängiger Mechanismus existiert, der über einen direkten C-C-Bindungsbruch an der 5-Position verläuft.^[59] Schon seit der Entdeckung von fdC wurde ein solcher, aus chemischer Sicht durchaus plausibler Mechanismus postuliert.^[5,60] *Pfaffeneder et al.*^[5] schlugen ursprünglich einen Mechanismus vor, der ausgehend des in wässriger Umgebung vorliegenden Hydrats von fdC über ein Vinylcarbanionintermediat verläuft. Später konnten *Schiesser et al.*^[51] zeigen, dass die Inkubation von synthetischen Oligonukleotiden, die isotopenmarkiertes cadC enthielten, mit Stammzellextrakt zu einer Decarboxylierung des schweren cadC zum isotopenmarkierten dC führte. Ein analoges Ergebnis wurde erhalten, wenn nur mit einer Kombination von Cystein und basisch wirkenden Aminosäuren inkubiert wurde, die das aktive Zentrum eines Enzyms simulieren sollten. Ein potentieller Mechanismus der enzymatischen Deformylierung von fdC ist in Abbildung 3-5 gezeigt.

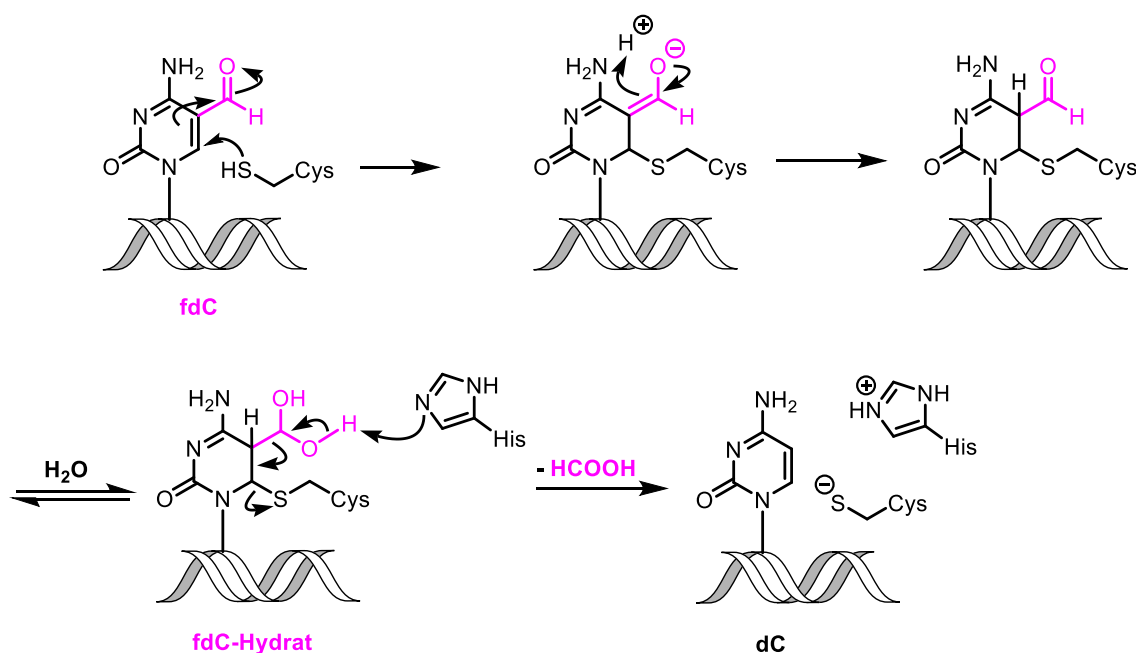


Abbildung 3-5: Postulierter Mechanismus der Deformylierung von genomischem fdC.

Ausgehend vom Hydrat des fdC erfolgt der nukleophile Angriff der enzymatischen Cystein-Seitenkette an die 6-Position des Pyrimidinrings. Über Enamin-Imin-Tautomerisierung kommt es zur Absättigung der C5-C6-Bindung. Daraufhin führt eine Deprotonierung der exozyklischen Hydrateinheit zur Abspaltung von Ameisensäure. Eine abschließende Eliminierung des Thio-Nukleophils unter Rearomatisierung des Pyrimidinrings liefert das deformylierte dC.

Während der oben dargestellte Mechanismus anhand der Erkenntnisse von *Pfaffeneder et al.*^[5] und *Schiesser et al.*^[51] nur ein Postulat darstellt, konnten *Iwan et al.*^[59] kürzlich beweisen, dass die aktive Demethylierung über das Intermediat des genomischen fdC tatsächlich über einen TDG-unabhängigen Mechanismus mittels direktem C-C-Bindungsbruch *in vivo* verlaufen kann. Dabei führten sie unter anderem Zellverfütterungsexperimente mit einem $^{13}\text{C}_5$ - $^{15}\text{N}_2$ -isotopenmarkierten fdC-Analogon durch und konnten unter Verwendung von hochsensitiver Massenspektrometrie in Kombination mit Ultrahochdruck-Flüssigchromatographie (UHPLC) zeigen, dass in der DNA das isotopenmarkierte Deformylierungsprodukt $^{13}\text{C}_5$ - $^{15}\text{N}_2$ -dC detektiert werden konnte.^[59] Des Weiteren setzten sie auch 2'-fluorierte Nukleosid-Derivate ein, die aufgrund ihrer veränderten elektronischen Eigenschaften nicht über BER durch TDG ausgeschnitten werden können.^[59,61] Der Fluorsubstituent dient des Weiteren auch der Detektion der Nukleosid-Derivate bei der anschließenden massenspektrometrischen Analyse. Bei Zellverfütterungsexperimenten mit 2'-Fluoro-fdC (F-fdC) konnte bei unterdrückter Reparatur durch TDG eine beträchtliche Deformylierungsaktivität hin zum 2'-Fluoro-dC (F-dC) *in vivo* nachgewiesen werden.^[59] Durch weitere Experimente, die durch Ausschaltung der TET-Enzyme keine weitere Oxidation der fdC-Derivate zum cadC erlaubten, konnte hier ebenfalls eine Decarboxylierung nach TET-katalysierter Oxidation zum cadC ausgeschlossen werden. Darüber hinaus zeigten massenspektrometrische Untersuchungen des in der Zelle gelösten monomeren Nukleosidpools, dass dort kein Deformylierungsprodukt der fdC-Derivate nachgewiesen werden konnte. Zusammenfassend konnte also festgestellt werden, dass eine direkte Deformylierung von genomischem fdC zum dC unter C-C-Bindungsspaltung *in vivo* stattfindet.^[59] Der exakte mechanistische Ablauf dieser Reaktion wurde jedoch bis zu diesem Zeitpunkt nicht untersucht.

3.4. Methoden zur Sequenzierung der epigenetisch relevanten Cytosin-Modifikationen

Neben den oben diskutierten, Massenspektrometrie-basierten Untersuchungen, die unabhängig von der DNA-Sequenz auf globalem Level erfolgten, ist es auch von großer Bedeutung für das Verständnis der Funktion dieser Modifikationen, ihr Auftreten im spezifischen DNA-Sequenzkontext zu untersuchen.

Die Sequenzierung von DNA, die neben den vier kanonischen Nukleosiden auch selektiv die modifizierten Cytidin-Nukleoside erfassen kann, ist seit Jahren Gegenstand intensiver Forschung.^[62–64] Die meisten Protokolle zur Sequenzierung dieser DNA-Modifikationen beruhen auf den Sequenziermethoden der ersten und zweiten Generation. Hier erfolgt nach Isolation der DNA und vor deren Sequenzierung eine Amplifikation des genetischen Materials via Polymerase-Kettenreaktion (*polymerase chain reaction*, PCR), um die Menge an zu sequenzierender DNA zu vervielfachen und damit das Signal bei der Sequenzierung der DNA zu verstärken.^[63] Die eigentliche Sequenzierung erfolgt dabei mittels bestehender Techniken wie beispielsweise der Illuminasequenzierung oder Pyrosequenzierung.^[65] Zur Unterscheidung der modifizierten Cytosin-Derivate, die sich bei einer PCR-Amplifikation alle wie unmodifiziertes Cytosin verhalten, werden die unterschiedlichen Reaktivitäten der Cytosin-Modifikationen mit dem Bisulfit-Reagenz ausgenutzt (siehe Abbildung 3-6).

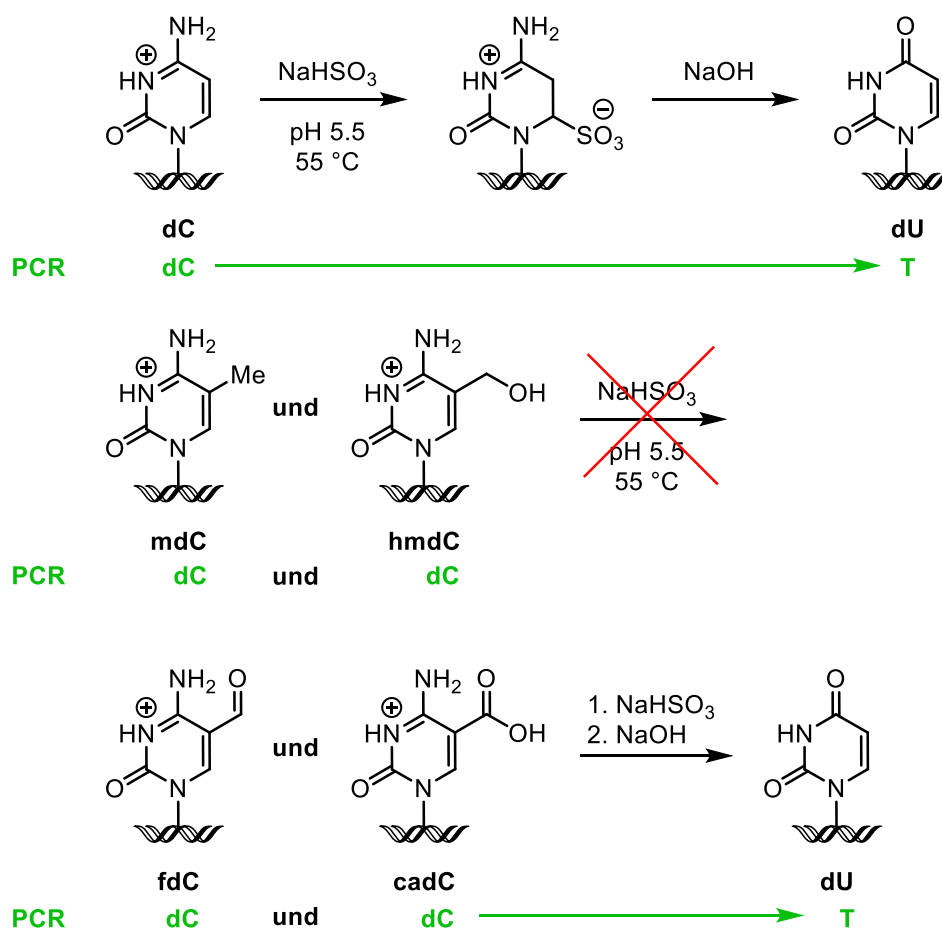


Abbildung 3-6: Darstellung der unterschiedlichen Reaktivitäten der Cytosin-Basen bei der Bisulfit-Behandlung und der daraus resultierenden Ergebnisse bei einer PCR-Amplifikation.

Die Behandlung von DNA mit Natriumhydrogensulfit in saurem pH führt im Falle von unmodifiziertem Cytosin zur Desaminierung der Nukleobase zum Uracil,^[66,67] das bei einer PCR-Amplifikation dann wie Thymin (T) und nicht mehr wie Cytosin behandelt wird. mdC und hmdC hingegen zeigen diese Reaktivität gegenüber dem Bisulfit-Reagenz nicht und werden deswegen nach der Bisulfit-Behandlung der DNA weiterhin wie Cytosin amplifiziert. Bei der Bisulfit-Behandlung von genomischem fdC und cadC kommt es neben der Desaminierung auch zusätzlich zur Deformylierung bzw. Decarboxylierung des jeweiligen Substituenten an der 5-Position, wodurch das Ablesemuster bei der PCR-Amplifikation wie beim unmodifizierten Cytosin zum Thymin verändert wird.^[68] Das ähnliche Reaktionsverhalten von dC, fdC und cadC sowie von mdC und hmdC bei der Bisulfit-Behandlung resultiert dabei in einem nicht eindeutig unterscheidbaren Sequenzierergebnis.

Um eine Unterscheidung dennoch zu ermöglichen, wurden diverse Protokolle entwickelt, die auf dem Vergleich mehrerer Sequenzierexperimente beruhen. Zur Sequenzierung von fdC kann dies beispielsweise mit einem milden Reduktionsmittel wie Natriumborhydrid zu hmdC reduziert werden, um so eine andere Reaktivität der Nukleobase bei der Bisulfit-Behandlung zu erzielen. Ein Vergleich mit den Sequenzierdaten einer unbehandelten Probe liefert schließlich die fdC-Positionen im Sequenzkontext.^[69] Eine andere Möglichkeit stellt die Derivatisierung von fdC mit Aldehyd-spezifischen Hydroxylamin-Reagenzien dar, was im Falle einer Reaktion von fdC mit Ethoxyamin zu einer Inversion der Reaktivität bei der Bisulfit-Behandlung führt und somit beim Vergleich mit einer unbehandelten Probe Aufschluss über die fdC-Positionen gibt.^[70] Ähnliche Methoden, die auf der chemischen Umsetzung einer Cytosin-Modifikation vor der Bisulfit-Amplifikation zur gezielten Veränderung ihrer Reaktivität beruhen, existieren auch für hmdC^[71] und cadC^[72]. Für fdC wurden darüber hinaus auch Bisulfit-freie Protokolle zur gezielten Veränderung des Verhaltens bei der PCR-Amplifikation entwickelt, die nichtsdestotrotz auf einem Vergleich mehrerer Sequenzierexperimente beruhen.^[73,74] Alles in allem verlangen die oben beschriebenen Protokolle zur selektiven Sequenzierung einer bestimmten Cytosin-Modifikation grundsätzlich mehrstufige Verfahren, da die epigenetische Information bei der notwendigen PCR-Reaktion verloren wird. Eine Unterscheidung von Nukleobasen mit ähnlichem Reaktionsverhalten wird dann nur durch den Vergleich mehrerer Sequenzierexperimente möglich (siehe Abbildung 3-7).

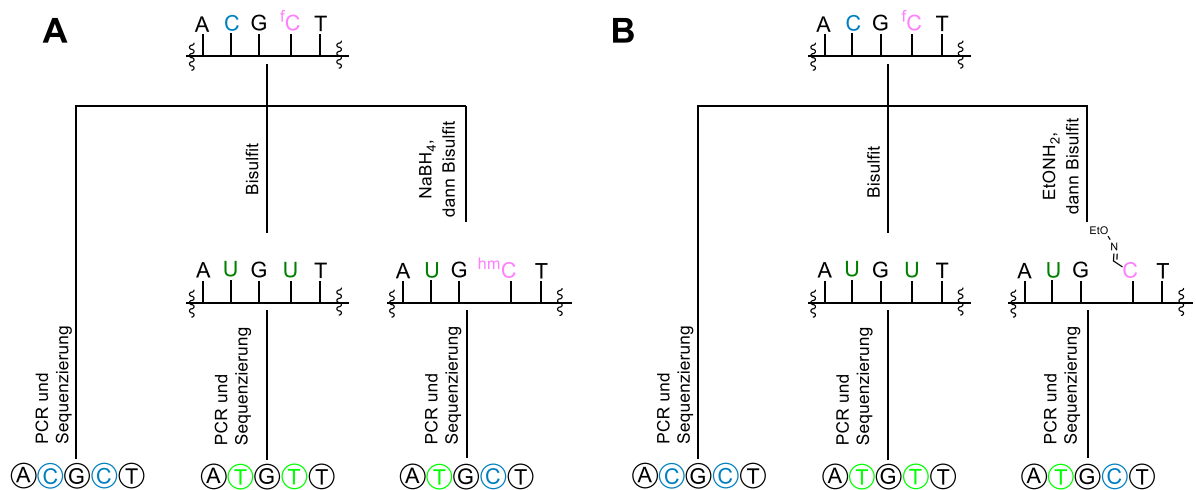


Abbildung 3-7: Darstellung von Methoden zur Sequenzierung von fdC durch Veränderung der Reaktivität bei der Bisulfitbehandlung durch A) Reduktion und B) chemische Derivatisierung.

Um ein möglicherweise fehleranfälligeres, mehrstufiges Vorgehen zu umgehen, bieten sich moderne Sequenziermethoden der dritten Generation an. Bei diesen wird die DNA auf Einzelmolekülebene sequenziert, was eine PCR-Amplifikation des zu sequenzierenden genetischen Materials und den damit einhergehenden Verlust der epigenetischen Information vermeidet. Bei der sogenannten *single molecule real time*-Sequenzierung (SMRT-Sequenzierung), die von der Firma *Pacific Biosciences* (*PacBio*) entwickelt wurde, wird das zu sequenzierende genetische Material spektroskopisch auf Einzelmolekülniveau sequenziert. Dazu wird jeweils ein DNA-Molekül durch eine in einer bestrahlten Kavität fixierten Polymerase repliziert, wobei für die Synthese des Gegenstranges Farbstoff-markierte Nukleotid-Derivate verwendet werden. Da das Licht für die Bestrahlung in der Kavität evaneszenten Charakter aufweist, d.h. dass dessen Intensität mit dem Abstand zum Eindringen in die Kavität exponentiell abfällt, wird selektiv der Einbau der Farbstoff-markierten Nukleotid-Derivate in den Komplementärstrang durch die fixierte Polymerase spektroskopisch beobachtet. Durch Verwendung von verschiedenen farbigen Markierungen kann so in Echtzeit auf Einzelmolekülniveau DNA sequenziert werden.^[75] Da diese SMRT-Sequenzierung für eine DNA-Probe parallel in mehreren 10.000 Kavitäten durchgeführt werden kann und die DNA vor der Sequenzierung durch die Ligation mit *Hairpin*-Primern bekannter Sequenz in zyklische Strukturen überführt wird, können lange Ableselängen mit gleichzeitig geringer Fehleranfälligkeit durch computerunterstützte Datenanalyse erhalten werden.^[75]

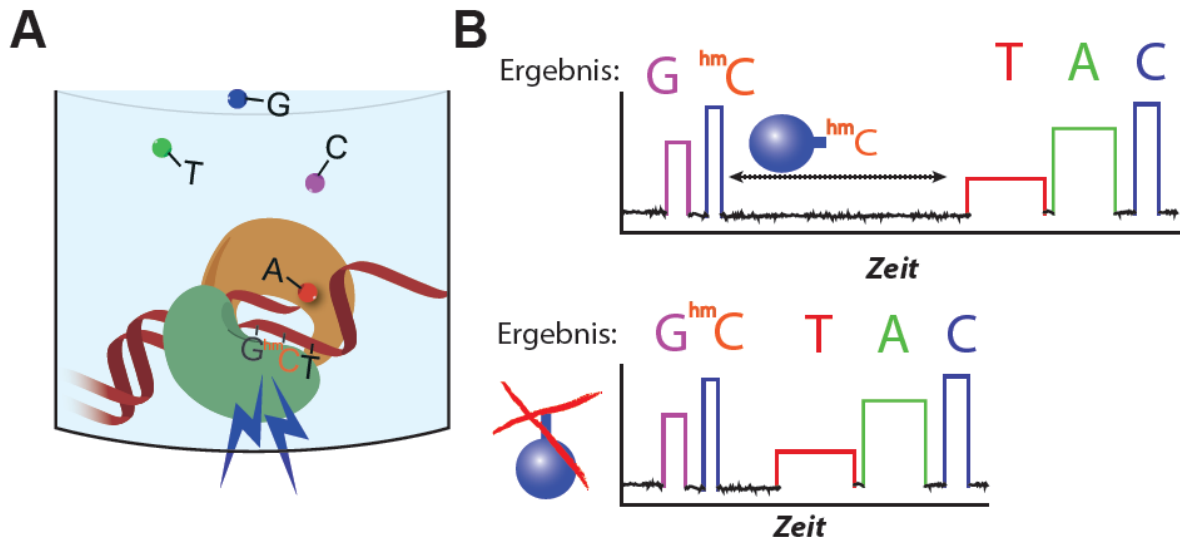


Abbildung 3-8: Schematische Darstellung einer Kavität der SMRT-Sequenzierung (A) und mögliche detektierte Signale mit (oben) und ohne Derivatisierung (unten) von hmdC (B).

Die Unterscheidung von kanonischem Cytosin und dessen modifizierten Derivaten kann beispielsweise durch die Zeit zwischen den detektierten Lichtsignalen (*interpulse duration*, IPD) beim Einbau komplementären Nukleotide erfolgen. So ist bekannt, dass vergleichsweise längere IPDs gemessen werden, falls ein Nukleotid komplementär zu einer Cytosin-Modifikation im Vergleich zum unmodifizierten Cytosin eingebaut wird. Allerdings sind diese Verlängerungen der IPD kurz.^[76] Durch selektive chemische Modifikation von hmdC wurde herausgefunden, dass beispielsweise eine enzymatische Glykosylierung der exozyklischen Hydroxymethylgruppe in signifikant längeren IPDs bei der Prozessierung des derivatisierten hmdCs resultiert.^[77,78] Dadurch können die modifizierten Nukleoside selektiv und sicher bei der bioinformatischen Datenanalyse erkannt und somit als solche in das Ergebnis des Sequenzierexperiments integriert werden (siehe Abbildung 3-8). Eine selektive Derivatisierung von genomischem fdC zu dessen verbesserter Detektion über verlängerte IPDs bei den Sequenzierungsverfahren der 3. Generation ist bisher noch nicht bekannt.

3.5. Beispiele von RNA-Modifikationen: Inosin und Queuosin

Im Gegensatz zur DNA, bei der nur wenige Modifikationen bekannt sind, weisen Ribonukleotide der RNA deutlich mehr Modifikationen auf.^[79,80] Die Vielzahl an verschiedenen Modifikationen sowie die hohe Modifikationsdichte innerhalb der RNA lässt sich durch deren komplexes Aufgabenspektrum erklären.^[80] RNA kann in kodierende RNA (*messenger* RNA, mRNA) und nicht-kodierende RNA (z.B. *transfer*

RNA, tRNA, *ribosomal* RNA, rRNA, *small-interfering* RNA, siRNA, *long non-coding* RNA, lncRNA, *micro* RNA, miRNA, etc.) unterschieden werden. Vielen Vertretern der kurzen, nicht-kodierenden RNA werden dabei insbesondere regulatorische Funktionen zugeordnet.^[81,82] Auch katalytische Aktivität von RNA in Form der sog. Ribozyme ist bekannt.^[83] Insgesamt sind bisher mehr als 170 verschiedene Modifikationen von Ribonukleosiden charakterisiert, wobei diese deutlich in der Komplexität variieren können.^[79,80] Dabei kann zwischen einfachen Modifikationen, die unter anderem Methylierungen, Acetylierungen oder Thiomethylierungen an vielen Positionen des Purin- oder Pyrimidinrings sowie an der Riboseeinheit umfassen,^[84] und Hypermodifikationen unterschieden werden (siehe Abbildung 3-9).^[48,85] Hypermodifikationen zeigen eine veränderte Struktur der Nukleobase oder vergleichsweise große, komplizierte Substituenten.^[48]

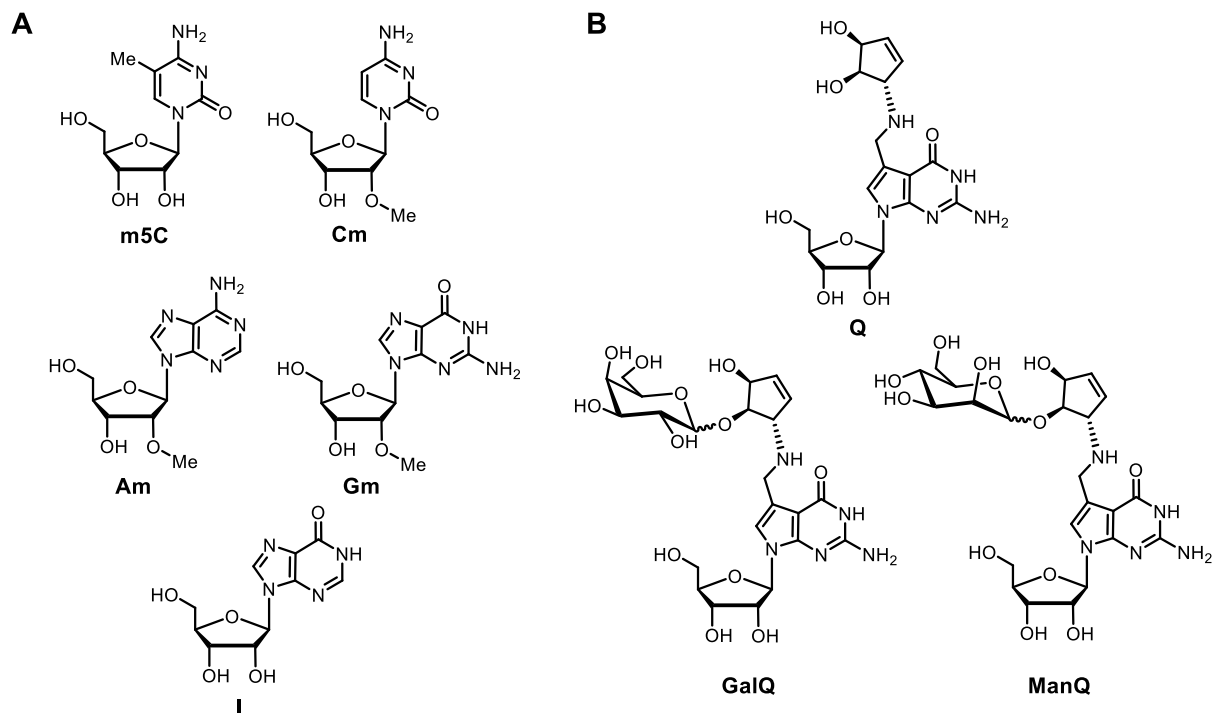


Abbildung 3-9: Beispiele für (A) einfache RNA-Modifikationen und (B) RNA-Hypermodifikationen (m5C: 5-Methylcytidin, Cm: 2'-Methylcytidin, Am: 2'-Methyladenosin, Gm: 2'-Methylguanosin, I: Inosin, Q: Queuosin, GalQ: Galaktosylqueuosin, ManQ: Mannosylqueuosin).

Queuosin (Q) ist ein besonders komplexer Vertreter der hypermodifizierten RNA Modifikationen, bei dem nicht nur die Nukleobase strukturell modifiziert ist.^[86–92] Diese liegt hier als 7-Deazaguanosinheterozyklus, bei dem der *N*7-Stickstoff durch einen C7-Kohlenstoff substituiert ist, vor. Die 7-Position weist darüber hinaus allerdings auch noch eine Verknüpfung zu einem Cyclopentendiol-Ringsystem auf. Diese besondere RNA Modifikation kann an der Wobble-Position 34 der tRNA-

Anticodonschleife von verschiedenen mitochondrialen und cytosolischen tRNAs (tRNA^{Asp}, tRNA^{Tyr}, tRNA^{Asn}, tRNA^{His}) in vielen Spezies inklusive einiger Eukaryoten gefunden werden.^[86–92] Sie dient dabei zur Interaktion mit entweder einem Cytidin (C) oder einem Uridin (U) im Codonbereich der mRNA. Des Weiteren kann die Q-Modifikation in einigen höheren Eukaryoten auch noch weiter modifiziert werden. Es ist bekannt, dass Q in der cytosolischen tRNA^{Tyr} und tRNA^{Asp} am Cyclopentendiol-Substituenten entweder mit Galaktose oder Mannose glykosyliert vorliegen kann.^[93,94] Die exakten stereo- und regiochemischen Verknüpfungen dieser Strukturen sowie deren definierte biologische Funktion sind bisher allerdings nicht exakt aufgeklärt.

Eine deutlich besser untersuchte und verstandene RNA-Modifikation stellt Inosin (I) dar. Diese kommt unter anderem in der mRNA vor und entsteht dort im Zuge des sog. *RNA editings*.^[95] In diesem Fall kann die 6-Position von Adenosin (A) in mRNA durch die *Adenosine Deaminases Acting on RNA* (ADARs) hydrolytisch desaminiert werden.^[96,97] Dies führt zu veränderten Basenpaarungseigenschaften, da I anders als A nicht mehr mit U sondern bevorzugt mit C die Basenpaarung ausbildet.^[98] Daraus können diverse Effekte, wie eine veränderte Genexpression, alternative Muster des Spleißens oder der Inkorporation anderer Aminosäuren, resultieren.^[99,100] Das *A-to-I-editing* stellt auf Grund dieser Effekte einen wichtigen Beitrag zur posttranskriptionellen Regulation dar, indem es die strukturelle und funktionelle Vielfalt der mRNA und damit auch schließlich die des Proteoms der Zellen erweitert.^[99] Da ein fehlerhaftes *A-to-I-RNA editing* im Menschen mit neurologischen oder psychischen Störungen sowie mit dem Auftreten von Krebserkrankungen in Zusammenhang gebracht werden kann^[98,101–103], wird dieser Effekt als essentiell betrachtet. Trotz des hohen Vorkommens von Inosin in der RNA sind die exakten strukturellen Veränderungen des *A-to-I-editings* auf die doppelsträngige mRNA sowie das exakte Ausmaß des *editings* bisher nicht vollständig aufgeklärt.

4. Aufgabenstellung

Im Rahmen der vorliegenden Dissertationsarbeit sollen mehrere Projekte bearbeitet werden, die allesamt zum Verständnis modifizierter Basen in DNA und RNA führen sollen.

Im ersten Teil der Arbeit sollen Studien im Zusammenhang mit der fdC-Modifikation in der DNA durchgeführt werden. Denn seitdem gezeigt werden konnte, dass die aktive Deformylierung von genomischem fdC über einen selektiven, direkten C-C-Bindungsbruch erfolgen kann,^[59] stellen sich besonders zwei Fragen: Einerseits ist bisher ungeklärt, wie der Mechanismus der direkten Deformylierung chemisch abläuft, andererseits ist es natürlich von besonders großem Interesse das die Deformylierung katalysierende Enzym zu identifizieren. Die mechanistischen Studien sollen auf Grundlage von Zellverfütterungsexperimenten mit neuen, synthetischen 6-Aza-2'-desoxycytidin-Analoga der epigenetisch relevanten Modifikationen durchgeführt werden (Kapitel 5.1 und 6.1), die es ermöglichen direkt den Einfluss eines nukleophilen Angriffs auf die 6-Position des Pyrimidinrings zu untersuchen (siehe Abbildung 4-1).

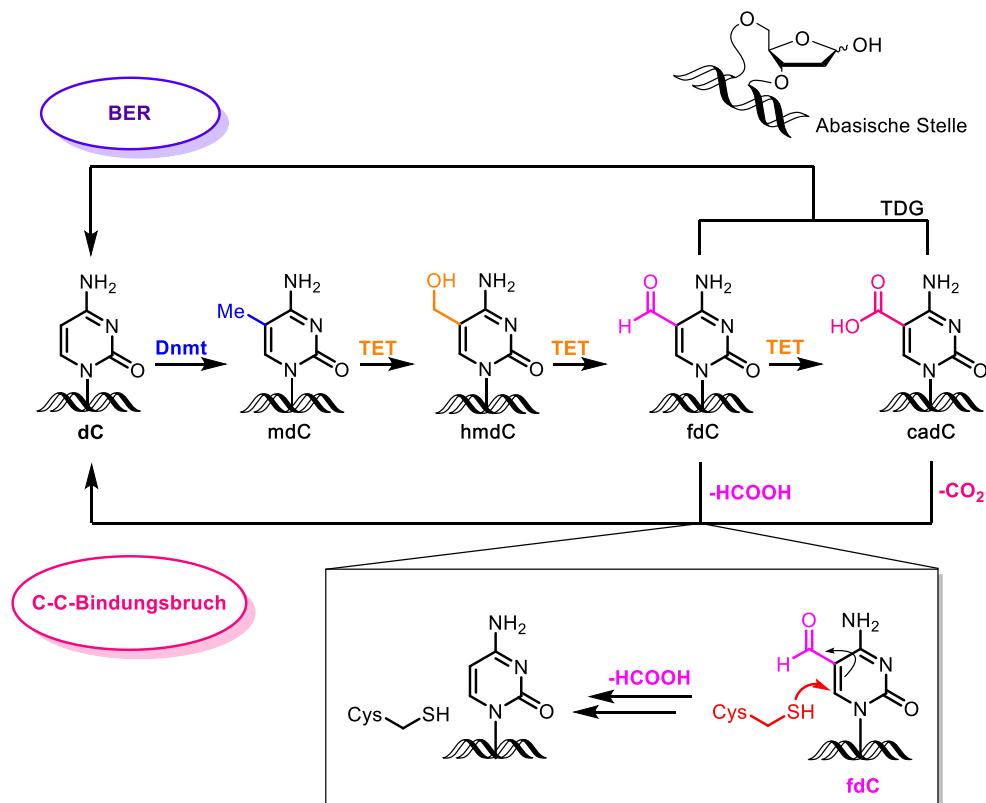


Abbildung 4-1: Einordnung des untersuchten Prozesses sowie des Mechanismus der aktiven, direkten DNA-Demethylierung

Außerdem sollen mit Hilfe eines spezifisch isoto­penmarkierten 2'-Desoxycytidins Informationen zur Quantifizierung einer direkten Deformylierung gewonnen werden (Kapitel 6.2). Im Hinblick auf die Untersuchung der beteiligten Enzyme rückten besonders die Sirtuine in den Vordergrund potentieller Kandidaten. Diese Hypothese sollte auf Basis enzymatischer *in vitro* Studien unter Verwendung synthetischer DNA-Oligonukleotide, die isoto­penmarkiertes fdC enthalten, untersucht werden (Kapitel 6.3).

Zum besseren Verständnis des strukturellen und dynamischen Einflusses der fdC-Modifikation auf den DNA-Doppelstrang sollen auch gezielte NMR-Untersuchungen kurzer synthetischer Oligonukleotide, die isoto­penmarkiertes oder nicht-markiertes fdC enthalten, in Kooperation mit der Arbeitsgruppe von *Dr. Petra Rovó* an der *LMU München* durchgeführt werden (Kapitel 5.2). Zur sequenzabhängigen Analyse von fdC sollen darüber hinaus Aldehyd-spezifische, sterisch anspruchsvolle Hydroxylaminreagenzien zur selektiven Derivatisierung von genomischem fdC entwickelt werden. Diese könnten bei einer Echtzeit-Sequenzierung von fdC zu einer Verlängerung der detektierten IPDs führen, um die Sequenzierung von fdC selektiv auf Einzelmolekülniveau zu ermöglichen (Abbildung 4-2, Kapitel 6.4).

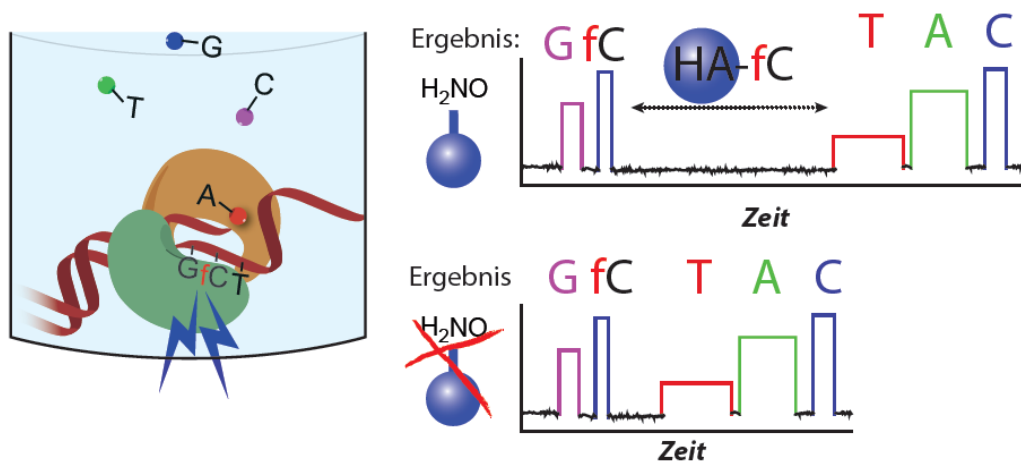


Abbildung 4-2: Prinzipdarstellung des Projekts zur SMRT-Sequenzierung von fdC.

Im Hinblick auf die medizinisch-chemische Kontrolle der Methylierung von dC zu mC soll in Analogie zu dem bereits bekannten carbazyklischen, hydrolysestabilen Analogon von Decitabin^[32] ebenfalls ein carbazyklisches Derivat von 5-Azacytidin synthetisiert und dessen Hydrolyseeigenschaften sowie dessen Einfluss auf die

Methylierung von DNA und Inkorporation in RNA untersucht werden (Abbildung 4-3, Kapitel 6.5).

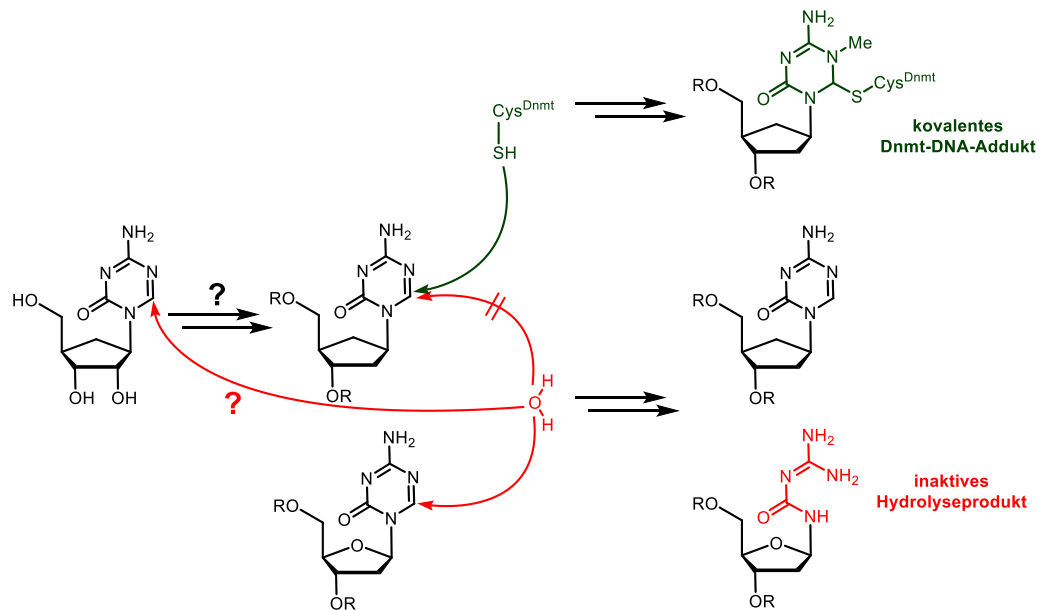


Abbildung 4-3: Schematische Vergleichsdarstellung des erwarteten Hydrolyseverhaltens von 5-Aza-C und dem carbazyklischen Analogon.

Im Rahmen der Untersuchungen im Hinblick auf Struktur und Eigenschaften von RNA-Modifikationen sollen Inosin sowie die glykosylierten Queuosine GalQ und ManQ genauer betrachtet werden.

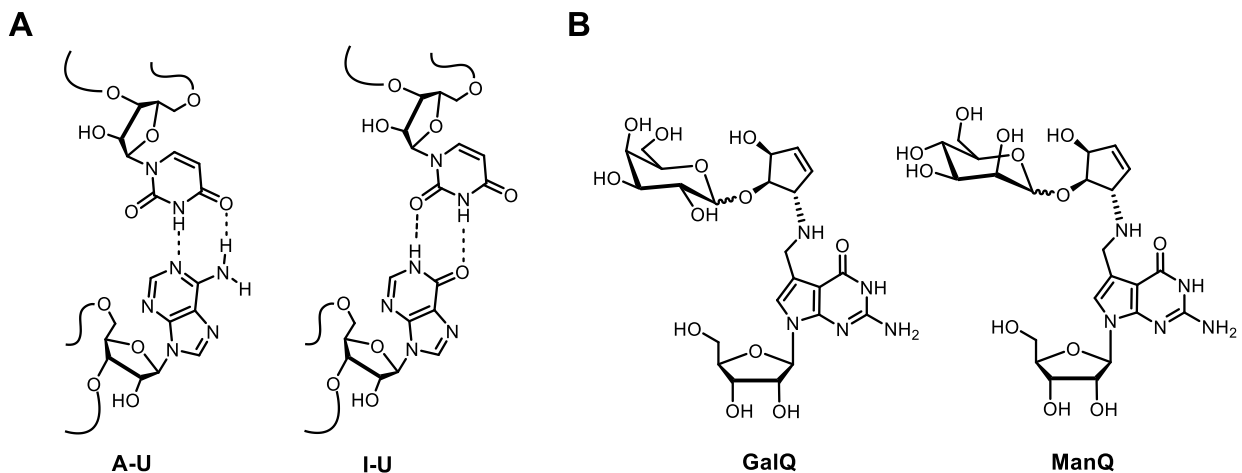


Abbildung 4-4: A) Basenpaarungsverhalten des A-U und postuliertes Basenpaarungsverhalten von I-U; B) Darstellung der RNA-Hypermodifikationen GalQ und ManQ.

Da bisher nicht geklärt ist, wie das A-zu-I-*editing* die Struktur doppelsträngiger RNA verändern kann und darüber hinaus welche Positionen des Inosins an der Basenpaarung beteiligt sind, sollen RNA Oligonukleotide eines RNA-Doppelstrangs synthetisiert werden, die Inosin mit gezielter Isotopenmarkierung enthalten sollen (Abbildung 4-4 A). Dieser RNA-Duplex soll dann im Rahmen einer Kooperation mit

der Arbeitsgruppe von *Prof. Michael Sattler* an der *TU München* NMR-spektroskopisch untersucht werden (Kapitel 6.6).

Da die exakten regio- und stereochemischen Verknüpfungen der Strukturen der RNA-Hypermodifikationen GalQ und ManQ bisher nicht exakt aufgeklärt sind, sollen diese Moleküle dafür auf Grundlage bestehender Arbeiten des Arbeitskreises *Carell* synthetisiert werden. In Koinjektionsexperimenten sollen diese dann mit den isolierten Naturstoffen verglichen werden, um so möglicherweise die Strukturen bestätigen zu können (Abbildung 4-4 B, Kapitel 5.3).

5. Veröffentlichte Arbeiten

5.1. Analysis of an active deformylation mechanism of 5-formyl-deoxycytidine (fdC) in stem cells

Alexander Schön[#], Ewelina Kaminska[#], Florian Schelter[#], Eveliina Ponkkonen, Eva Korytiaková, Sarah Schiffers and Thomas Carell*

([#] geteilte Erstautorenschaft)

Prolog

Nachdem von *Iwan et al.*^[59] gezeigt werden konnte, dass eine direkte Deformylierung von genomischem fdC existiert, stellt sich nun natürlich die Frage, wie der Prozess *in vivo* mechanistisch ablaufen kann. Basierend auf Studien von *Schiesser et al.*^[51] ist bekannt, dass eine Kombination von Cystein und basisch wirkenden Aminosäuren wie Histidin oder Arginin zur Decarboxylierung von cadC in synthetischen Oligonukleotiden führen kann. Dies legt den Schluss nahe, dass zur Deformylierung von fdC analog zur Decarboxylierung von cadC basische und nukleophile Seitenketten im aktiven Zentrum des oder der an diesen Prozessen beteiligten Enzyme enthalten sein müssten. Der nukleophile Angriff der Thiol-Funktionalität wird hierbei, analog zum Mechanismus der DNA-Methyltransferasen, an die 6-Position des Pyrimidinrings der modifizierten Nukleoside postuliert.

Im Rahmen dieser Arbeit wurden die Synthesen von 6-Aza-fdC (a-fdC), welches an der 6-Position ein Stickstoffatom trägt und in der Konsequenz keinen nukleophilen Angriff an dieser Position zulässt, und dessen potentiellen Deformylierungsprodukts 6-Aza-dC (a-dC) entwickelt. Eine Parallelverfütterung von a-fdC zusammen mit dem bekannten Molekül F-fdC an verschiedene Zelllinien zeigte, dass *in vivo* für F-fdC starke, wohingegen für a-fdC keine Deformylierungsaktivitäten detektiert werden konnten. Dies legt nahe, dass ein nukleophiler Angriff an die 6-Position von fdC zur *in vivo*-Deformylierung notwendig ist.

Autorenbeiträge

Im Rahmen dieses Projekts entwickelte ich die Synthesen von a-fdC und a-dC und führte diese durch. Ich synthetisierte außerdem das potentielle Desaminierungsprodukt von a-fdC 6-Aza-5-formyl-2'-desoxyuridin (a-fdU), welches bei den

massenspektrometrischen Untersuchungen der Verfütterungsexperimente nicht nachgewiesen werden konnte. Außerdem untersuchte ich die Derivatisierung von a-fdC, fdC und F-fdC mit Hydroxylaminen und entwickelte daraus die Derivatisierungsstrategie mit Methoxyamin, die schließlich zur erfolgreichen Detektion von a-fdC bei der massenspektrometrischen Analyse der verdauten DNA-Extrakte führte. *Ewelina Kaminska* führte alle zellbiologischen Verfütterungsexperimente durch. *Florian Schelter* entwickelte massenspektrometrische Methoden zur Detektion der zu untersuchenden Moleküle und führte DNA-Verdaus sowie Messungen durch. *Eveliina Ponkkonen* führte die Studien zur Untersuchung der Reaktivitäten von fdC, F-fdC und a-fdC mit Bisulfit durch. *Eva Korytiaková* synthetisierte F-fdC. *Dr. Sarah Schiffers* entwickelte ebenfalls Methoden zur Detektion der 6-Aza-Nukleoside. Alle Daten wurden zusammen von den Erstautoren zusammen mit *Thomas Carell* diskutiert und interpretiert.

Anmerkung

Da die von mir entwickelten Vorschriften zur Synthese der Verbindungen sowie zur Derivatisierung von a-fdC in der *Supporting Information* des Artikels dargestellt sind, ist diese im Folgenden auch abgedruckt.

Lizenz

Kopie der *Open Access* Publikation auf Basis der *Creative Commons Attribution-NonCommercial License (CC BY-NC 4.0)*, die den nicht-kommerziellen Abdruck des Artikels erlaubt.

Artikel einsehbar unter:

<https://onlinelibrary.wiley.com/doi/full/10.1002/anie.202000414>

Lizenz einsehbar unter:

<https://creativecommons.org/licenses/by-nc/4.0/>

Epigenetics

International Edition: DOI: 10.1002/anie.202000414
German Edition: DOI: 10.1002/ange.202000414

Analysis of an Active Deformylation Mechanism of 5-Formyl-deoxycytidine (fdC) in Stem Cells

Alexander Schön⁺, Ewelina Kaminska⁺, Florian Schelter⁺, Eveliina Ponkkonen, Eva Korytiaková, Sarah Schiffrers, and Thomas Carell^{*}

Dedicated to Dr. Klaus Römer on the occasion of his 80th birthday

Abstract: The removal of 5-methyl-deoxycytidine (mdC) from promoter elements is associated with reactivation of the silenced corresponding genes. It takes place through an active demethylation process involving the oxidation of mdC to 5-hydroxymethyl-deoxycytidine (hmdC) and further on to 5-formyl-deoxycytidine (fdC) and 5-carboxy-deoxycytidine (cadC) with the help of α -ketoglutarate-dependent Tet oxygenases. The next step can occur through the action of a glycosylase (TDG), which cleaves fdC out of the genome for replacement by dC. A second pathway is proposed to involve C–C bond cleavage that converts fdC directly into dC. A 6-aza-5-formyl-deoxycytidine (a-fdC) probe molecule was synthesized and fed to various somatic cell lines and induced mouse embryonic stem cells, together with a 2'-fluorinated fdC analogue (F-fdC). While deformylation of F-fdC was clearly observed *in vivo*, it did not occur with a-fdC, thus suggesting that the C–C bond-cleaving deformylation is initiated by nucleophilic activation.

The nucleobase modification 5-formyl-deoxycytidine (fdC, **1**) is found in stem cells during early development and in the brain.^[1–5] These tissues are particularly rich in 5-hydroxymethyl-deoxycytidine (hmdC) from which fdC (**1**) is produced.^[6,7] The formation of hmdC and fdC requires oxidation reactions that are performed by α -ketoglutarate-dependent Tet enzymes, with 5-methyl-deoxycytidine (mdC) being the initial starting molecule.^[8–10] This cascade of oxidation reactions is a part of an active demethylation process, in which mdC as a silencer of transcription is replaced by unmodified dC.^[11] The central molecule that is removed seems to be

fdC.^[12,13] It can be cleaved out of the genome by a dedicated DNA glycosylase, which creates an abasic site that is further processed, leading to the insertion of an unmodified dC.^[14] Because abasic sites are harmful DNA-repair intermediates that can cause genome instability, it was suggested early on that fdC might be directly deformylated to dC by C–C bond cleavage.^[15,16] Evidence for the existence of such a direct deformylation process was recently reported.^[17] Model studies showed that direct deformylation of fdC and potentially also decarboxylation of 5-carboxy-deoxycytidine (cadC) are indeed possible.^[15] Nevertheless, it requires activation of the nucleobase by a nucleophilic addition to the C6 position. For fdC, an additional hydrate formation on the formyl group seems to be necessary, as depicted in Figure 1 A. Although activation with a helper nucleophile is well known as the central mechanistic process during the methylation of dC to mdC by DNA methyltransferases (Dnmts),^[18,19] it remains to be confirmed whether such activation occurs *in vivo* as well.

[*] M. Sc. A. Schön,^[+] E. Kaminska,^[+] M. Sc. F. Schelter,^[+] M. Sc. E. Ponkkonen, M. Sc. E. Korytiaková, Dr. S. Schiffrers, Prof. Dr. T. Carell
Department of Chemistry, Ludwig-Maximilians Universität München
Butenandtstr. 5–13, 81377 München (Germany)
E-mail: thomas.carell@lmu.de
Homepage: <https://www.carellgroup.de>

[+] These authors contributed equally to this work.

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under:
<https://doi.org/10.1002/anie.202000414>.

© 2020 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited, and is not used for commercial purposes.

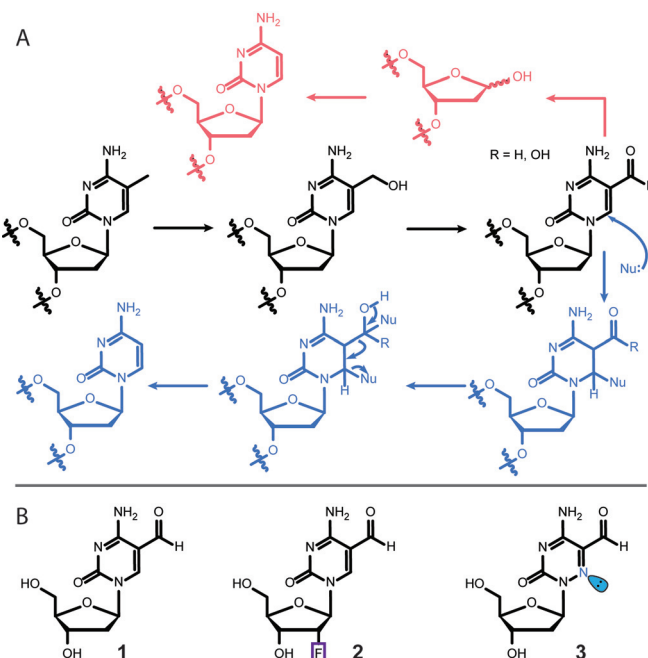


Figure 1. A) The mdC removal pathways that involve oxidation to hmdC, fdC, and cadC followed by either base-excision repair (magenta) or C–C bond cleavage (blue). B) Structures of fdC (**1**) and the two probe molecules **2** and **3** used for this study.

In this work, we investigated this hypothesis with two probe molecules, 2'-fluorinated-fdC (F-fdC, **2**) and 6-aza-fdC (a-fdC, **3**). The two compounds were simultaneously fed to different cell types, including primed stem cells. This led to random incorporation of these bases at the "C" sites in the respective genomes. Furthermore, it led to the presence of F-fdC and a-fdC not only at CpG sites. Ultrasensitive UHPLC-QQQ-MS² was subsequently used to interrogate the chemical processes that occur at F-fdC and a-fdC in the genomes. The data show that while F-fdC is efficiently deformylated, this does not occur for a-fdC. The only difference between the two nucleobases is the presence of an in-ring nitrogen atom (6-aza atom), which features a lone pair that prohibits nucleophilic addition. These results thus provide strong evidence that nucleophilic activation is the central governing mechanistic event that is required for C–C bond cleavage *in vivo*.

The fluorinated nucleoside F-fdC (**2**) was recently introduced by us as a deformylation probe.^[17] Compound **2** is an antimetabolite that is effectively incorporated into the genomes of growing cells. The 2'-fluoro group is required to block all types of glycosylases, so that base-excision repair is efficiently inhibited. This ensures high levels of F-fdC (**2**) in the genome, as required to observe potential deformylation processes.

The synthesis of the novel nucleoside a-fdC (**3**) is depicted in Scheme 1. The synthesis was started with bromo pyruvic acid (**4**), which we first converted into the semicarbazone **5**, followed by conversion into the acid chloride, subsequent cyclization, and hydrolysis to give hydroxymethylated 6-azauracil (**6**).^[20] Vorbrüggen nucleosidation with Hoffer's chlorosugar subsequently provided the nucleoside **7** as a mixture of the α - and β -anomers, which could be separated by recrystal-

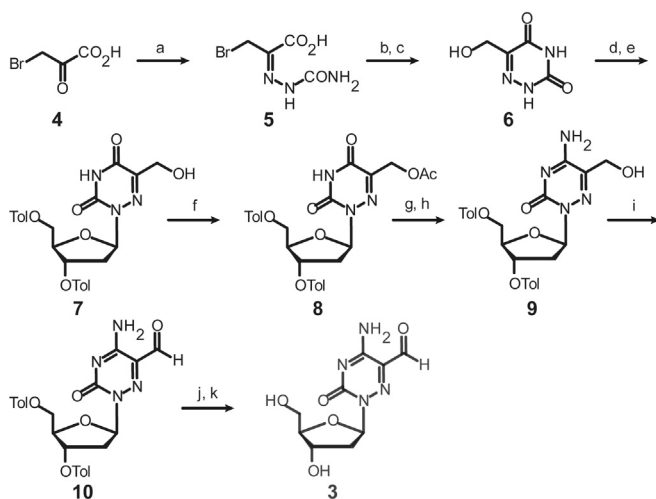
lization. Next, we acetyl-protected the hydroxymethyl group to give **8**, and then used a standard procedure to convert the U base **8** into the C-derived base **9** by amination of the 4-triazole intermediate with ammonium hydroxide. This led to the concomitant cleavage of the acetyl protecting group. Dess–Martin oxidation of **9** to **10** and final removal of the toluoyl groups furnished the 6-aza-5-formyl-deoxycytidine nucleoside (a-fdC) **3** in a good total yield of 22 % with respect to **6** (Supporting Information).

Compound **3** features a nitrogen atom instead of a carbon atom at the 6-position, which possesses a lone pair that blocks any nucleophilic addition to this position. Compound **3** is consequently a perfect model system to investigate whether such a nucleophilic activation is required for the deformylation, as mechanistically postulated (Figure 1 A).

The nucleosides **2** and **3** were subsequently added at a concentration of 350 μM to the media of Neuro-2a, RBL-2H3, CHO-K1 cells for 72 hours (see the Supporting Information). During this time, the nucleosides are converted *in vivo* into the corresponding triphosphates and then incorporated into the genome of the dividing cells. Initial studies in which we fed the nucleosides individually allowed us to determine that neither compound decreases cell viability up to a concentration of 400 μM , thus the experiments were conducted below the toxicity level. In addition, we tested **2** and **3** at 350 μM on E14 TDG +/– and –/– mouse embryonic stem cells (mESCs) under a three-day priming process with C/ R media. This system allowed us to exclude the BER pathway, leading to a detectable and quantifiable accumulation of natural fdC (see the Supporting Information). After three days, the cells were harvested and lysed, and the genomic DNA was extracted using an optimised protocol (see the Supporting Information). This was followed by an enzymatic digestion of gDNA to single nucleosides and analysed according to a method that we reported recently in detail.^[21]

The obtained nucleoside mixture containing mostly the canonical nucleosides dA, dC, dG, and dT, plus the non-canonical nucleosides mdC, hmdC, and fdC, as well as the incorporated molecules F-fdC and a-fdC and their potential downstream products (F-dC, F-mdC, a-dC, a-mdC). Nucleosides were separated by ultra-HPL chromatography and characterized by coupling of the UHPLC system to a triple-quadrupole mass spectrometer. For exact quantification of the nucleosides by isotope dilution, isotopically labelled standards of F-fdC and of the product F-dC were spiked into the analysis mixture as internal standards (see the Supporting Information). To enable exact quantification, calibration curves using these standards were determined (see the Supporting Information). Quantification was performed in the linear region.

During the analysis, we noted that an unusually low amount of a-fdC (**3**) was detected because it showed a broad elution profile with very low intensity (Figure 2B). All attempts to sharpen the elution profile in order to gain sensitivity failed. NMR analysis of compound **3** showed the reason for broad elution profile (see the Supporting Information). Due to the additional electron-withdrawing in-ring nitrogen atom, compound **3** exists partially as its hydrate in aqueous solution (20%, see the Supporting Information).



Scheme 1. Synthesis of the probe molecule a-fdC (**3**). a) semicarbazide-HCl, NaOAc, HOAc, 0 °C to r.t., 2.5 h, 49%. b) pyridine, SOCl₂, 80 °C, 75 min. c) H₂O, 110 °C, 17 h, 74 % over 2 steps. d) TMSCl, HMDS, 135 °C, 75 min, then e) Hoffer's chlorosugar, CHCl₃, r.t., 17 h, 56 % over 2 steps. f) Ac₂O, pyridine, r.t., 22 h, 96%. g) 1,2,4-triazole, POCl₃, NEt₃, MeCN, 0 °C to r.t., 18 h, then h) NH₄OH, 1,4-dioxane, 40 °C, 5 h, 84%. i) Dess–Martin periodinane, CH₂Cl₂, –15 °C to r.t., 1 h, 89%. j) NaOMe, MeOH, benzene, r.t., 1.5 h, then k) reversed-phase HPLC, 54 %.

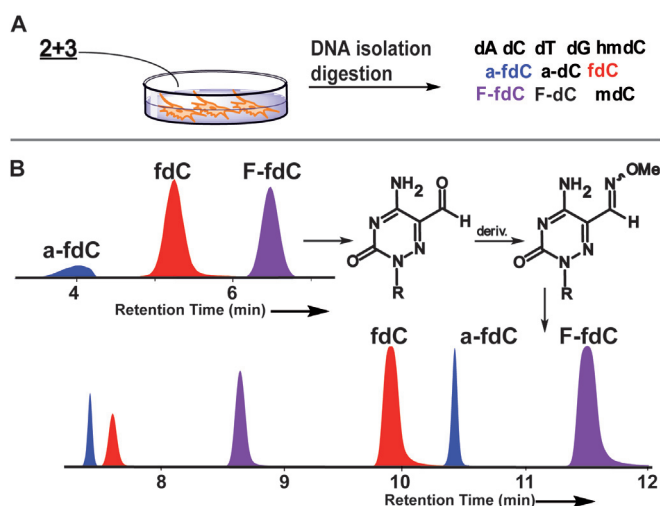


Figure 2. A) Overview of the experimental steps with the feeding and analysis. B) Analysis scheme and the reaction of a-fdC with methoxyamine to block hydrate formation and of a typical UHPL-chromatogram before (C-8 column) and after derivatization (C-18 column) for exact quantification. Peak splitting is due to isomerization (blue peaks: a-fdC, red peaks: fdC, and purple peaks: F-fdC).

Although the ease of hydrate formation may foster deformylation, the hydrate/carbonyl equilibrium makes efficient detection of compound **3** basically impossible. In order to circumvent the problem, we started to derivatize a-fdC (**3**) before analysis with methoxyamine. Addition of CH_3ONH_2 (150 mM) to the digestion solution indeed provided the methoxyoxime of a-fdC in quantitative yield after just 15 min at 25 °C and pH 10. The naturally present fdC (**1**) and the probe molecule F-fdC (**2**), however, react as well, but unfortunately not quantitatively. To reduce impurities during MS measurements, we decided against using a catalyst for oxime formation. We therefore decided to analyse the digested DNA in two batches. The first one contained the digested untreated DNA to quantify all bases other than a-fdC. In the second batch, we treated the digested DNA with methoxyamine for a-fdC quantification. For quantification of the derivatized a-fdC, we constructed an external calibration curve (see the Supporting Information).

With this method in hand, we next quantified all nucleosides present in the genome of the cells treated with a mixture of **2** + **3**.

Figure 3A shows that we indeed detected the fluorinated F-dC (**2**), thus confirming very efficient deformylation activity. We tested different cell types and found different levels of deformylation activity. But in all cases, the conversion of F-fdC into F-dC was clearly detectable. Most interesting is that we observed the highest deformylation activity in cells associated with neuronal properties. This is in line with neurons featuring the highest levels of hmdC and fdC. In contrast, Figure 3B shows that for a-fdC (**3**), we were unable to detect any formation of the deformylated compound a-dC despite the high propensity of **3** to exist in the hydrated form, which is one prerequisite for efficient C–C bond cleavage. This result suggests that the ability to react

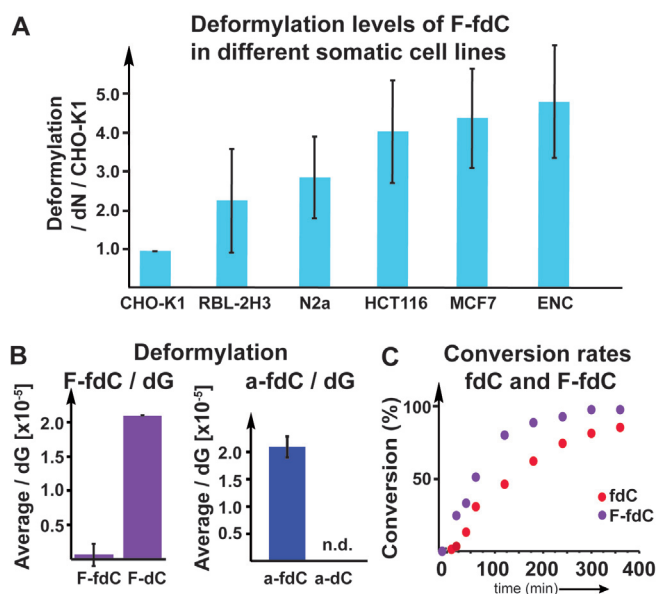


Figure 3. A) Deformylation data for F-fdC in different cell types, showing that F-fdC is deformylated in very different cells. Deformylation rate was calculated by the $\text{F-dC} + \text{F-mdC}/\text{dN}$ per $\text{F-fdC}/\text{dN}$, then the values were normalized to the cell line with the lowest deformylation level (CHO-K1 = 1). B) The deformylation of F-fdC/dG and a-fdC/dG, showing the induced differences due to C6-carbon-to-nitrogen exchange. C) The bisulfite data show that the deformylation of fdC and F-fdC is comparable, thus showing that the 2'-F substitution has only a small accelerating effect, whereas the reaction of a-fdC could not be detected.

with a nucleophile at the 6-position is also required in vivo for efficient deformylation.

In order to substantiate this result, we next performed in vitro studies with bisulfite. Bisulfite is a strong nucleophile that has been reported to cause deformylation of fdC by first attacking the C6 position, followed by conversion of the C5,C6-saturated fdC adduct into the bisulfite adduct, which then undergoes deformylation.^[22] The deformylated product dC is then further converted into dU by the well-known bisulfite-induced deamination reaction of dC (see the Supporting Information). Indeed, when we reacted fdC with bisulfite, we observed efficient deformylation and deamination to dU. We then studied to what extent the reaction is influenced by the 2'-F atom present in F-fdC, in order to estimate whether the in vivo deformylation could be just the result of the 2'-F atom. Treatment of F-fdC with bisulfite also led to deformylation and deamination to F-dU, and indeed the reaction is a little faster compared to fdC (see Figure 3C). Although the difference is measurable, it is in total rather small. With these data in hand, we can conclude that we may overestimate the amount of deformylation that can occur with fdC lacking the 2'-F atom. We can certainly exclude that deformylation in vivo occurs only with F-fdC. It is unfortunate that we are unable to measure the direct deformylation of fdC because of the presence of efficient BER processes. A TDG $-/-$ cell line showed a huge increase in fdC compared to the TDG $+/-$, whereas a-fdC and F-fdC stay constant, thus showing that these compounds are indeed not repaired by the

TDG protein (see the Supporting Information). The bisulfite studies, however, show that the F-fdC compound is not a perfect but sufficient reporter of this C–C bond cleavage. Treatment of a-fdC (**3**) with bisulfite did not provide the deformed product a-dC under any circumstances, showing that the inability to react with a nucleophile at the 6-position totally blocks the C–C bond cleavage. We can therefore conclude that the deformatation of fdC during active demethylation requires oxidation of mdC to fdC. fdC can undergo a direct C–C bond cleavage to dC, but this reaction requires a helper nucleophile to attack the C6-position, which is blocked in the case of a-fdC by the lone pair introduced by the C6-carbon-to-nitrogen exchange. While the chemistry that allows the transformation of fdC into dC is now elucidated, we next need to find the nucleophiles that perform the reaction in vivo.

Acknowledgements

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) GRK2338 (Project ID 321812289), SFB1309 (PID 325871075), SFB1361 (PID 393547839) and SPP1784 (PID 255344185). This project has received additional funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement n° EPIr 741912) and through a H2020 Marie Skłodowska-Curie Action (LightDyNAMics, 765866).

Conflict of interest

The authors declare no conflict of interest.

Keywords: demethylation · DNA modifications · epigenetics · formylcytidine

How to cite: *Angew. Chem. Int. Ed.* **2020**, *59*, 5591–5594
Angew. Chem. **2020**, *132*, 5639–5643

- [1] M. Wagner, J. Steinbacher, T. F. J. Kraus, S. Michalakakis, B. Hackner, T. Pfaffeneder, A. Perera, M. Müller, A. Giese, H. A. Kretzschmar, T. Carell, *Angew. Chem. Int. Ed.* **2015**, *54*, 12511–12514; *Angew. Chem.* **2015**, *127*, 12691–12695.
- [2] S. Liu, J. Wang, Y. Su, C. Guerrero, Y. Zeng, D. Mitra, P. J. Brooks, D. E. Fisher, H. Song, Y. Wang, *Nucleic Acids Res.* **2013**, *41*, 6421–6429.
- [3] C. X. Song, K. E. Szulwach, Q. Dai, Y. Fu, S. Q. Mao, L. Lin, C. Street, Y. Li, M. Poidevin, H. Wu, J. Gao, P. Liu, L. Li, G. L. Xu, P. Jin, C. He, *Cell* **2013**, *153*, 678–691.
- [4] T. Pfaffeneder, B. Hackner, M. Truss, M. Münzel, M. Müller, C. A. Deiml, C. Hagemeier, T. Carell, *Angew. Chem. Int. Ed.* **2011**, *50*, 7008–7012; *Angew. Chem.* **2011**, *123*, 7146–7150.
- [5] M. Bachman, S. Uribe-Lewis, X. Yang, H. E. Burgess, M. Iurlaro, W. Reik, A. Murrell, S. Balasubramanian, *Nat. Chem. Biol.* **2015**, *11*, 555–557.
- [6] S. Kriaucionis, N. Heintz, *Science* **2009**, *324*, 929–930.
- [7] D. Q. Shi, I. Ali, J. Tang, W. C. Yang, *Front. Genet.* **2017**, *8*, 100.
- [8] M. Münzel, D. Globisch, T. Carell, *Angew. Chem. Int. Ed.* **2011**, *50*, 6460–6468; *Angew. Chem.* **2011**, *123*, 6588–6596.
- [9] C. G. Spruijt, F. Gnerlich, A. H. Smits, T. Pfaffeneder, P. W. T. C. Jansen, C. Bauer, M. Münzel, M. Wagner, M. Müller, F. Khan, H. C. Eberl, A. Mensinga, A. B. Brinkman, K. Lephikov, U. Müller, J. Walter, R. Boelens, H. Van Ingen, H. Leonhardt, T. Carell, M. Vermeulen, *Cell* **2013**, *152*, 1146–1159.
- [10] S. Ito, L. Shen, Q. Dai, S. C. Wu, L. B. Collins, J. A. Swenberg, C. He, Y. Zhang, *Science* **2011**, *333*, 1300–1303.
- [11] F. Neri, D. Incarnato, A. Krepelova, S. Rapelli, F. Anselmi, C. Parlato, C. Medana, F. DalBello, S. Oliviero, *Cell Rep.* **2015**, *10*, 674–683.
- [12] M. Su, A. Kirchner, S. Stazzoni, M. Müller, M. Wagner, A. Schröder, T. Carell, *Angew. Chem. Int. Ed.* **2016**, *55*, 11797–11800; *Angew. Chem.* **2016**, *128*, 11974–11978.
- [13] T. Fu, L. Liu, Q. L. Yang, Y. Wang, P. Xu, L. Zhang, S. Liu, Q. Dai, Q. Ji, G. L. Xu, C. He, C. Luo, L. Zhang, *Chem. Sci.* **2019**, *10*, 7407–7417.
- [14] R. Rahimoff, O. Kosmatchev, A. Kirchner, T. Pfaffeneder, F. Spada, V. Brantl, M. Müller, T. Carell, *J. Am. Chem. Soc.* **2017**, *139*, 10359–10364.
- [15] S. Schiesser, T. Pfaffeneder, K. Sadeghian, B. Hackner, B. Steigenberger, A. S. Schröder, J. Steinbacher, G. Kashiwazaki, G. Höfner, K. T. Wanner, C. Ochsenfeld, T. Carell, *J. Am. Chem. Soc.* **2013**, *135*, 14593–14599.
- [16] S. Schiesser, B. Hackner, T. Pfaffeneder, M. Müller, C. Hagemeier, M. Truss, T. Carell, *Angew. Chem. Int. Ed.* **2012**, *51*, 6516–6520; *Angew. Chem.* **2012**, *124*, 6622–6626.
- [17] K. Iwan, R. Rahimoff, A. Kirchner, F. Spada, A. S. Schröder, O. Kosmatchev, S. Ferizaj, J. Steinbacher, E. Parsa, M. Müller, T. Carell, *Nat. Chem. Biol.* **2018**, *14*, 72–78.
- [18] A. Jeltsch, *ChemBioChem* **2002**, *3*, 274–293.
- [19] Q. Du, Z. Wang, V. L. Schramm, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 2916–2921.
- [20] I. V. Alekseeva, A. S. Shalamai, V. S. Shalamai, V. P. Chernetski, *Ukr. Khim. Zh. (Ukr. Ed.)* **1976**, *42*, 398–401.
- [21] F. R. Traube, S. Schiffrers, K. Iwan, S. Kellner, F. Spada, M. Müller, T. Carell, *Nat. Protoc.* **2019**, *14*, 283–312.
- [22] E. Kriukienė, Z. Liutkevičiūtė, S. Klimašauskas, *Chem. Soc. Rev.* **2012**, *41*, 6916–6930.

Manuscript received: January 9, 2020

Accepted manuscript online: January 30, 2020

Version of record online: February 25, 2020

Supporting Information

Analysis of an Active Deformylation Mechanism of 5-Formyl-deoxycytidine (fdC) in Stem Cells

*Alexander Schön⁺, Ewelina Kaminska⁺, Florian Schelter⁺, Eveliina Ponkkonen, Eva Korytiaková, Sarah Schiffers, and Thomas Carell**

anie_202000414_sm_miscellaneous_information.pdf

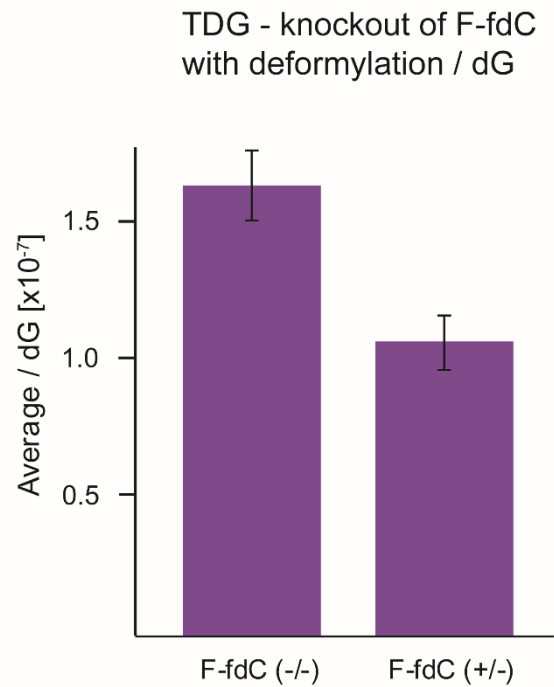
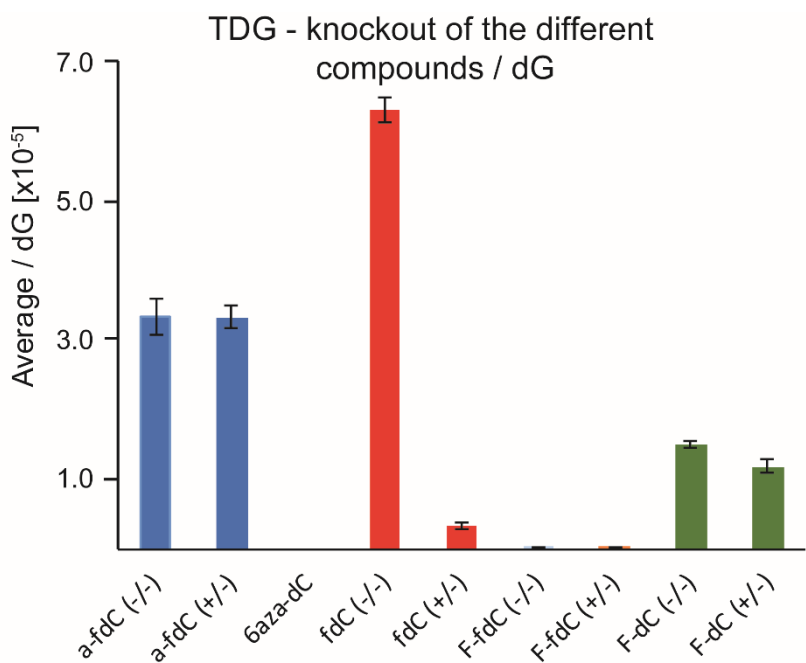
Supporting information

Table of contents

Supplementary figures	2
General Methods	6
Synthetic Procedures	7
Reactions of 1 , 2 , 3 and 15 with Methoxyamine	18
Reaction of 1 , 2 and 3 with NaHSO ₃	19
Cell culture	19
Chemical labelling and spiking	20
Analysis of labelled and unlabelled products of F-fdC and a-fdC	20
Digestion of the DNA of N2a cell line	22
Derivatization of a-fdC in digested gDNA of N2a cell line	23
Calibration curve for external quantification of a-fdC in gDNA of N2a cell lines	23
Quantification of concentration of the fed compounds	24
NMR spectra of synthesized compounds	26
Bibliography	38

Supplementary figures

Table 1: TDG knockout experiments of the fed nucleosides



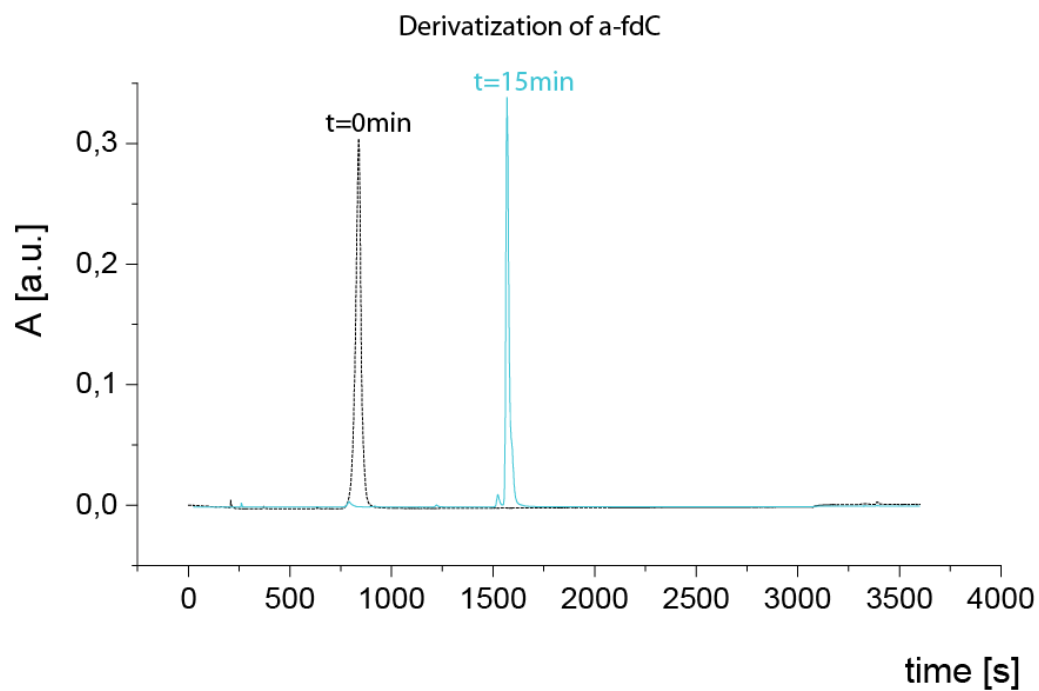


Figure 1: Quantitative derivatization of a-fdC with MeONH₂ after 15 min.

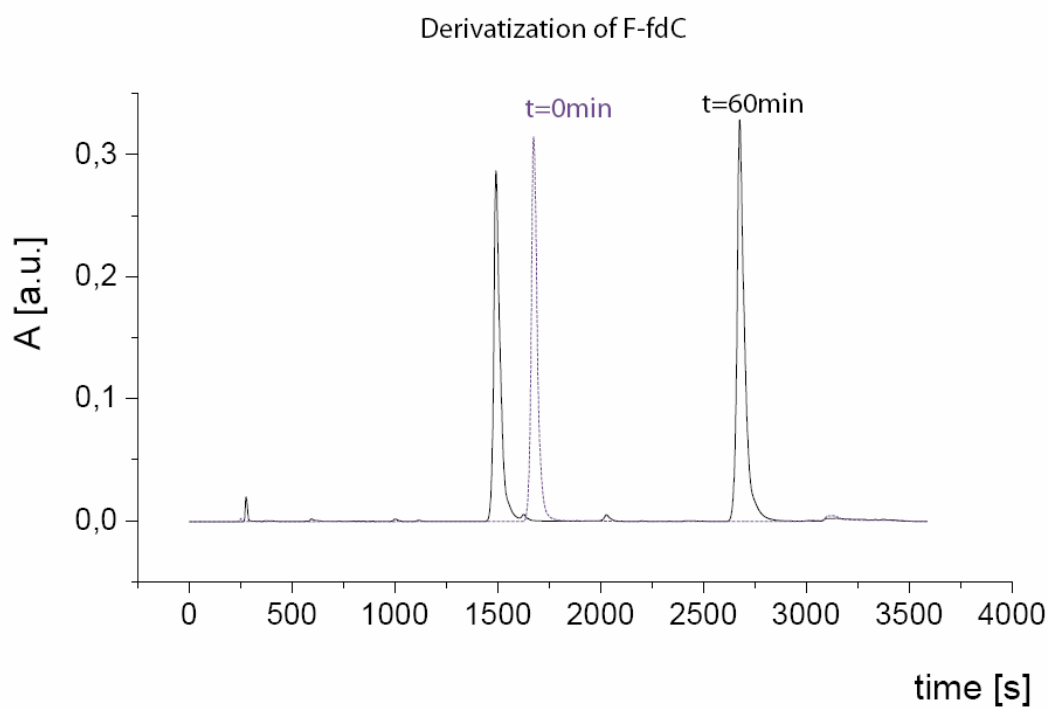


Figure 2: Derivatisation of F-fdC after 60 min, mixture of product and starting material.

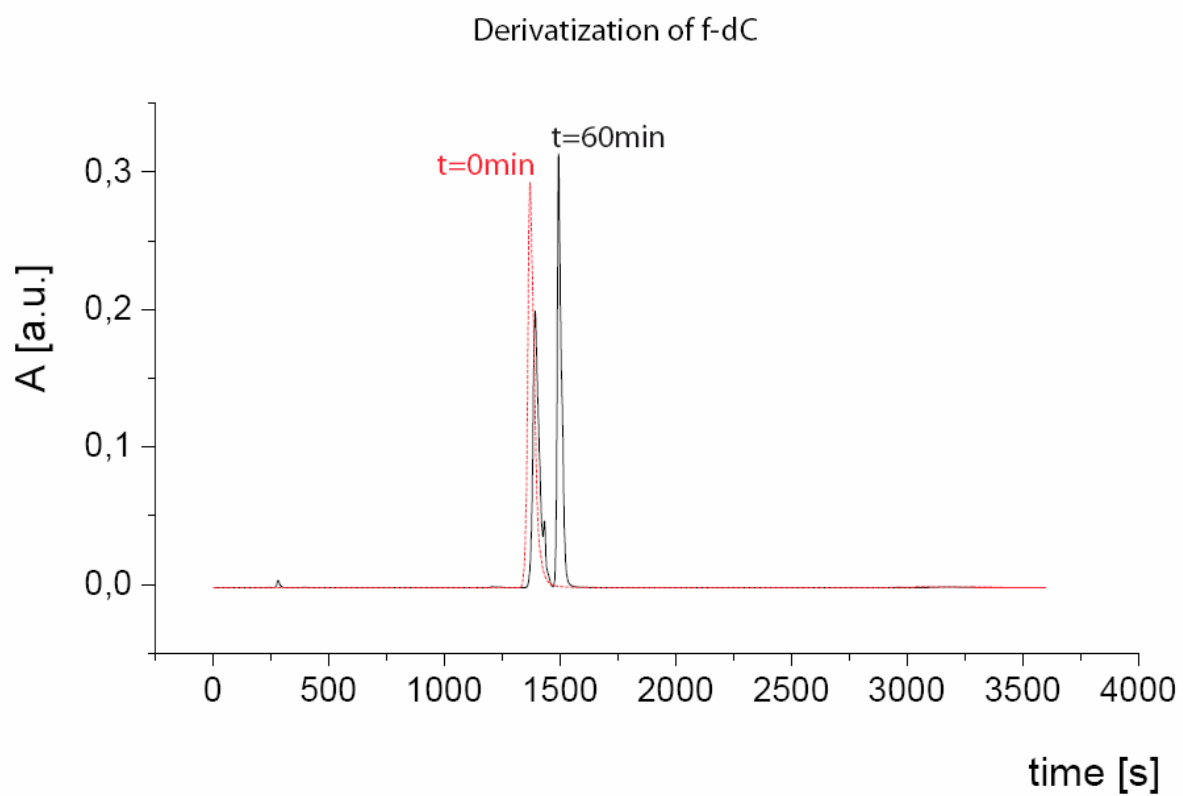


Figure 3: Derivatisation of fdC after 60 min, mixture of product and starting material.

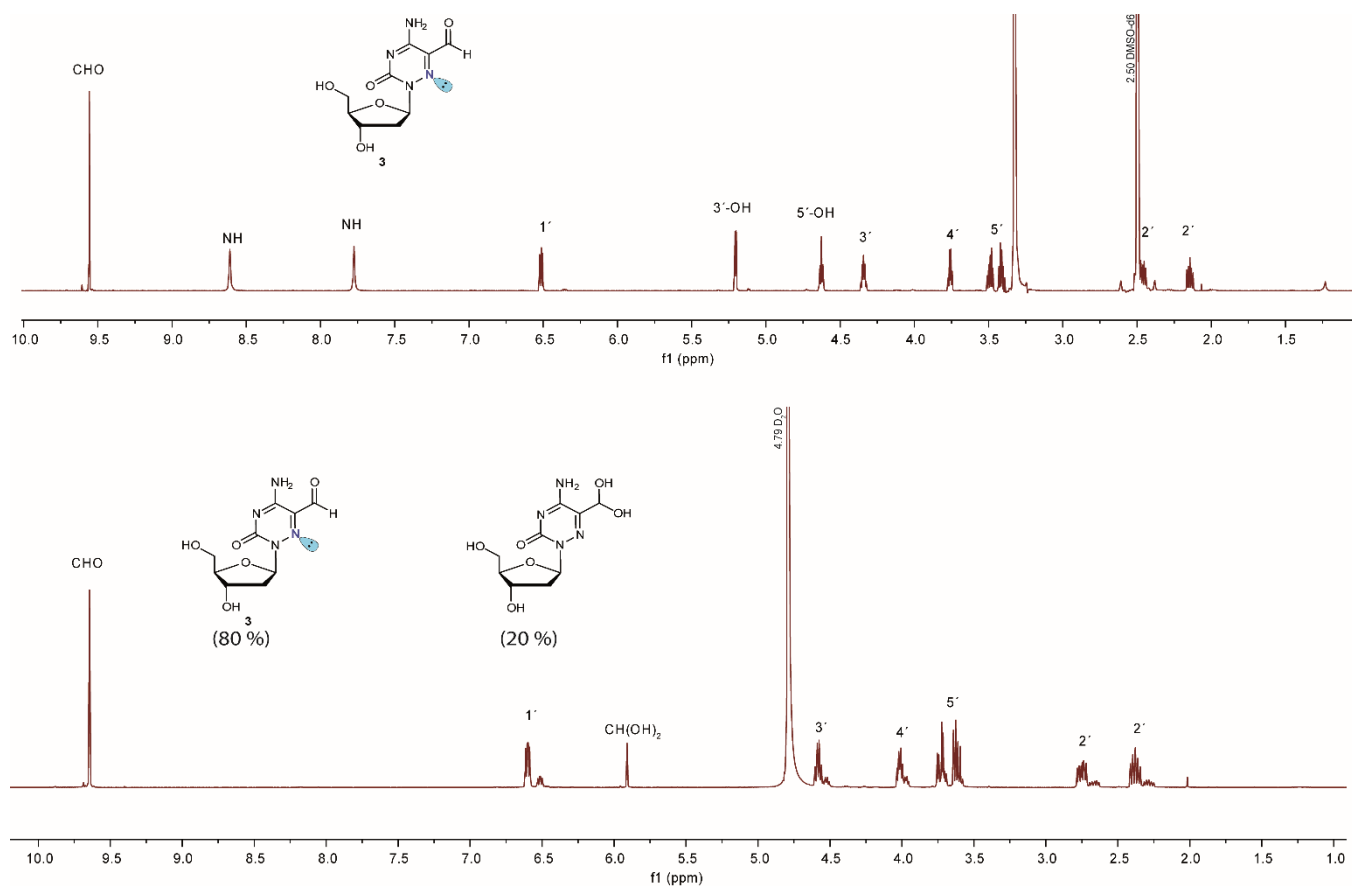


Figure 4: NMR spectra of α -fdC showing the hydrate formation of α -fdC in D_2O . Upper panel: NMR in $DMSO-d_6$. Lower panel: NMR in D_2O .

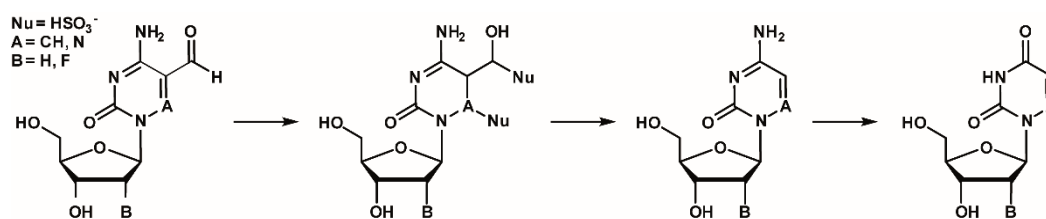


Figure 5: Proposed reaction mechanism with bisulfite

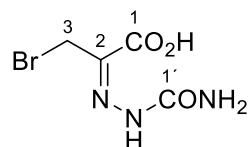
General Methods

Chemicals were purchased from *Sigma-Aldrich*, *TCI*, *Fluka*, *ABCR*, *Carbosynth* or *Acros Organics* and used without further purification. Solvents of reagent grade were purified by distillation. Reactions and column chromatography fractions were monitored by thin-layer chromatography (TLC) on silica gel F254TLC plates from *Merck KGaA*. Flash column chromatography was performed on *Geduran®* Si60 (40-63 µm) silica gel from *Merck KGaA* applying slight nitrogen pressure. Chemical transformations were conducted under nitrogen or argon atmosphere in oven-dried glassware unless otherwise specified. NMR spectra were recorded on *Bruker AVIIIHD 400* (400 MHz) or *Varian NMR-System600* (600 MHz) spectrometers. ¹H-NMR shifts were calibrated to the residual solvent resonances: CDCl₃ (7.26 ppm), DMSO-d₆ (2.50 ppm) and D₂O (4.79 ppm). ¹³C-NMR shifts were calibrated to the residual solvent: CDCl₃ (77.16 ppm), DMSO-d₆ (39.52 ppm). All NMR spectra were analysed using the program *MestReNova* 10.0.1 from *Mestrelab Research S. L.* Low resolution mass spectra were measured on a LT Q FT-ICR by *Thermo Finnigan GmbH*. High resolution mass spectra were measured by the analytical section of the Department of Chemistry of the Ludwigs-Maximilians-Universität München on a MAT 90 (ESI) from *Thermo Finnigan GmbH*. IR spectra were recorded on a *PerkinElmer Spectrum BX II FT-IR* system. Substances were applied as a film or directly as solids on the ATR unit. Analytical RP-HPLC was performed on an analytical HPLC *Waters Alliance* (2695 Separation Module, 2996 Photodiode Array Detector) equipped with the column *Nucleosil 120-2 C18* from *Macherey Nagel* applying eluent flow of 0.5 mL/min. Preparative RP-HPLC was performed on a HPLC *Waters Breeze* (2487 Dual λArray Detector, 1525 Binary HPLC Pump) equipped with the column VP 250/32 C18 from *Macherey Nagel*. A flowrate of 5 mL/min was applied.

Synthetic Procedures

Synthesis of 6-Aza-5-formyl-2'-deoxycytidine **3**

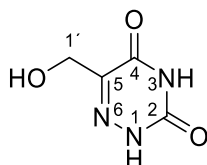
3-Bromopyruvic acid semicarbazone **5**



A solution of semicarbazide hydrochloride (3.34 g, 30.0 mmol, 1.0 eq.) and NaOAc (3.20 g, 38.6 mmol, 1.3 eq.) in 25.0 mL ddH₂O was added to a solution of 5.00 g 3-bromopyruvic acid (30.0 mmol, 1.0 eq.) in 15.0 mL glacial HOAc and 5.0 mL ddH₂O at 0 °C. The mixture was allowed to stand at room temperature for 2.5 h and at 0 °C for 19 h. The resulting white precipitate was filtered, washed with ice cold ddH₂O and Et₂O and dried subsequently under high vacuum to yield 3.26 g of the semicarbazone **5** as a mixture of (E)- and (Z)-isomers (14.6 mmol, 49 %) as a colorless solid.

¹H-NMR (400 MHz, DMSO-*d*₆): δ/ppm = 10.58 (s, 0.6H, NNH), 10.52 (s, 0.4H, NNH), 7.39 (br s, 0.8H, CONH₂), 6.92 (br s, 1.3H, CONH₂), 4.57 (s, 1.3H, 3-H), 4.46 (s, 0.8H, 3-H). **¹³C-NMR** (101 MHz, DMSO-*d*₆): δ/ppm = 163.51 (1-C), 155.63 (1'-C), 132.48 (2-C), 32.48 (3-C). **IR (ATR):** ν (cm⁻¹) = 3462 (w), 3245 (w), 2394 (br w), 1894 (br w), 1693 (m), 1445 (m), 1421 (s), 1364 (m), 1297 (m), 1234 (w), 1201 (s), 1164 (m), 1143 (m), 992 (m), 821 (s), 695 (m), 658 (s).

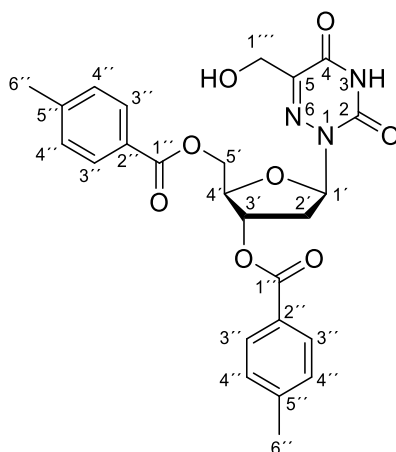
5-Hydroxymethyl-6-azauracil **6**



Compound **6** was synthesized according to a modified procedure of *Alekseeva et al.*^[1] Under Ar atmosphere 1.40 g semicarbazone **5** (6.2 mmol, 1.0 eq.) were suspended in 50.0 mL SOCl₂ and two drops of pyridine were added. The reaction mixture was refluxed at 80 °C for 75 min, allowed to cool to room temperature, filtered through Celite and concentrated *in vacuo* to approximately 15 mL. Crystallization at 0 °C for 4 d resulted in a yellow precipitate which was filtered, dissolved in 30 ml dry CH₂Cl₂ and concentrated to dryness. After drying under high vacuum the obtained yellow solid was suspended in 20.0 mL ddH₂O and refluxed at 110 °C for 17 h. Subsequently, the solution was concentrated *in vacuo*, the residue was dissolved in 15.0 mL ddH₂O and lyophilized to yield 657 mg 5-hydroxy-6-azauracil **6** (4.60 mmol, 74 %) as a beige solid.

¹H-NMR (400 MHz, DMSO-*d*₆): δ/ppm = 12.16 (s, 1H), 11.93 (s, 1H), 4.25 (s, 2H, 1'-H). **¹³C-NMR** (101 MHz, DMSO-*d*₆): δ/ppm = 156.75 (4-C), 149.55 (2-C), 144.33 (5-C), 58.02 (1'-C). **HRMS (ESI):** calculated for C₄H₄N₃O₃⁻ [M-H]⁻ 142.0258; found: 142.0258. **IR (ATR):** ν (cm⁻¹) = 3429 (w), 3197 (w), 3034 (w), 2846 (w), 1682 (s), 1476 (m), 1446 (m), 1415 (m), 1244 (m), 1207 (m), 1061 (m), 1030 (m), 869 (m), 811 (m), 742 (s).

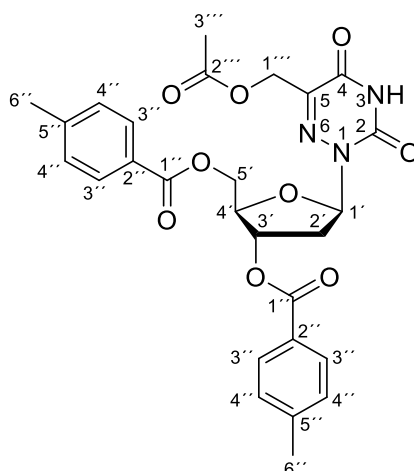
5-Hydroxymethyl-(3',5'-di-O-p-toluoyl)-6-aza-2'-deoxyuridine 7



In a heat-dried round bottom flask and under argon atmosphere 290 mg 5-hydroxymethyl-6-azauracil **6** (2.03 mmol, 1.0 eq.) were suspended in 10.64 mL hexamethyldisilazane and 351 μ L TMSCl were added. The mixture was refluxed at 135 °C for 75 min, the resulting brown solution was cooled to room temperature and concentrated *in vacuo* at 30 °C. The oily residue was dried under high vacuum for 1 h and subsequently dissolved in 6.38 mL dry CHCl_3 . 867 mg Hoffer's chlorosugar (2.23 mmol, 1.1 eq.) were added and the reaction mixture was stirred at room temperature for 17 h. The reaction mixture was poured into 250 mL 0.2 M aq. HCl and extracted with EtOAc (3 x 200 mL). Combined organic layers were dried over Na_2SO_4 and concentrated to dryness. Purification via silica gel column chromatography (iHex:EtOAc 3:1 \rightarrow iHex:EtOAc 1:1 \rightarrow EtOAc) yielded a 9:1 mixture of diastereomeric β - and α - nucleosidation products (760 mg, 1.53 mmol, 75 %) as a colorless solid. Recrystallization from absolute EtOH yielded 558 mg of the pure β -nucleoside **7** (1.13 mmol, 56%) as a colorless solid.

$^1\text{H-NMR}$ (600 MHz, CDCl_3): δ /ppm = 9.52 (s, 1H, 3-H), 7.96 – 7.88 (m, 4H, 3''-H), 7.26 – 7.19 (m, 4H, 4''-H), 6.73 (dd, J = 6.7, 6.6 Hz, 1H, 1'-H), 5.65 (ddd, J = 6.3, 3.1, 3.0 Hz, 1H, 3'-H), 4.73 (dd, J = 11.5, 7.0 Hz, 1H, 5'-H), 4.61 – 4.46 (m, 4H, 4'-H, 5'-H, 1'''-H), 2.99 (ddd, J = 14.3, 6.7, 6.5 Hz, 1H, 2'-H), 2.50 (ddd, J = 14.2, 6.8, 3.3 Hz, 1H, 2'-H), 2.42 (s, 3H, 6''-H), 2.38 (s, 3H, 6''-H). **$^{13}\text{C-NMR}$** (151 MHz, CDCl_3): δ /ppm = 166.70 (1''-C), 166.19 (1''-C), 155.59 (4-C), 148.47 (2-C), 145.58 (5-C), 144.54 (5''-C), 144.29 (5''-C), 129.94 (3''-C), 129.37 (4''-C), 129.30 (4''-C), 126.85 (2''-C), 126.59 (2''-C), 86.27 (1'-C), 82.74 (4'-C), 75.20 (3'-C), 64.36 (5'-C), 60.29 (1'''-C), 35.04 (2'-C), 21.87 (6''-C), 21.82 (6''-C). **HRMS (ESI $^+$)**: calculated for $\text{C}_{25}\text{H}_{26}\text{N}_3\text{O}_8^+$ $[\text{M}+\text{H}]^+$ 496.1714, found: 496.1715; calculated for $\text{C}_{25}\text{H}_{29}\text{N}_4\text{O}_8^+$ $[\text{M}+\text{NH}_4]^+$ 513.1980, found: 513.1975; calculated for $\text{C}_{25}\text{H}_{25}\text{N}_3\text{O}_8\text{Na}^+$ $[\text{M}+\text{Na}]^+$ 518.1534, found: 518.1529. **HRMS (ESI $^-$)**: calculated for $\text{C}_{25}\text{H}_{24}\text{N}_3\text{O}_8^-$ $[\text{M}-\text{H}]^-$ 494.1566; found: 494.1572. **IR (ATR)**: ν (cm^{-1}) = 1714 (s), 1611 (m), 1450 (w), 1273 (s), 1178 (m), 1105 (s), 1020 (w), 753 (m). **R_f** (iHex: EtOAc 1:1): 0.27.

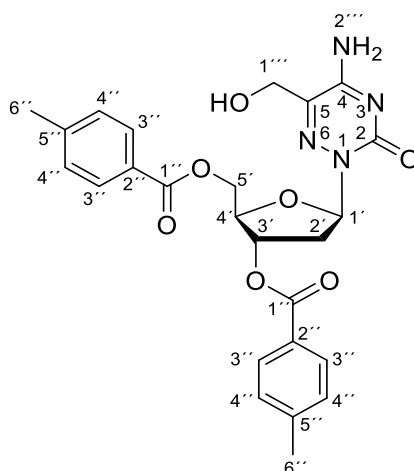
5-Methylacetate-(3',5'-di-O-p-toluoyl)-6-aza-2'-deoxyuridine **8**



273 mg hydroxymethyl nucleoside **7** (0.55 mmol, 1.0 eq.) were dissolved in 3.30 mL dry pyridine and 220 μ L acetic anhydride (2.33 mmol, 4.2 eq.) were added under argon atmosphere. The reaction mixture was stirred at room temperature for 22 h and stopped by addition of 5.0 mL absolute EtOH. The solution was concentrated *in vacuo*, coevaporated twice from dry toluene (2 x 15 mL) and the residue was purified by silica gel column chromatography ($\text{CH}_2\text{Cl}_2 \rightarrow \text{CH}_2\text{Cl}_2:\text{MeOH}$ 20:1) to obtain 283 mg of acetyl protected nucleoside **8** (0.53 mmol, 96 %) as a colorless foam.

$^1\text{H-NMR}$ (600 MHz, CDCl_3): δ/ppm = 8.54 (s, 1H, 3-H), 7.94 – 7.88 (m, 4H, 3''-H), 7.26 – 7.18 (m, 4H, 4''-H), 6.67 (dd, J = 6.9, 5.4 Hz, 1H, 1'-H), 5.69 (ddd, J = 7.4, 4.5, 3.2 Hz, 1H, 3'-H), 5.04 (d, J = 14.0 Hz, 1H, 1'''-H), 4.91 (d, J = 14.0 Hz, 1H, 1'''-H), 4.58 – 4.49 (m, 3H, 4'-H, 5'-H), 2.96 (ddd, J = 14.1, 6.8, 5.4 Hz, 1H, 2'-H), 2.51 (ddd, J = 14.1, 7.0, 4.5 Hz, 1H, 2'-H), 2.43 (s, 3H, 6''-H), 2.39 (s, 3H, 6''-H), 2.15 (s, 3H, 3'''-H). **$^{13}\text{C-NMR}$** (151 MHz, CDCl_3): δ/ppm = 170.33 (2'''-C), 166.33 (1''-C), 166.07 (1''-C), 154.46 (4-C), 147.79 (2-C), 144.52 (5''-C), 144.13 (5''-C), 141.45 (5-C), 129.94 (3''-C), 129.89 (3''-C), 129.38 (4''-C), 129.27 (4''-C), 127.05 (2''-C), 126.65 (2''-C), 86.29 (1'-C), 82.54 (4'-C), 74.86 (3'-C), 64.33 (5'-C), 59.91 (1'''-C), 35.38 (2'-C), 21.88 (6''-C), 21.82 (6''-C), 20.73 (3'''-C). **HRMS (ESI⁺)**: calculated for $\text{C}_{27}\text{H}_{31}\text{N}_4\text{O}_9$ $[\text{M}+\text{NH}_4]^+$ 555.2086, found: 555.2019. **HRMS (ESI⁻)**: calculated for $\text{C}_{27}\text{H}_{26}\text{N}_3\text{O}_9$ $[\text{M}-\text{H}]^-$ 536.1675; found: 536.1677. **IR (ATR)**: ν (cm^{-1}) = 1952 (w), 1711 (s), 1610 (m), 1442 (m), 1380 (m), 1269 (s), 1177 (m), 1100 (s), 1019 (m), 839 (w), 751 (s). **R_f** ($\text{CH}_2\text{Cl}_2:\text{MeOH}$ 20:1): 0.52.

5-Hydroxymethyl-(3',5'-di-O-p-toluoyl)-6-aza-2'-deoxycytidine **9**

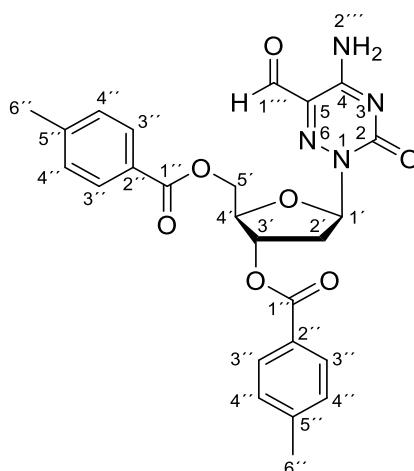


In a heat dried Schlenk-flask 487 mg 1,2,4-triazole (7.05 mmol, 9.0 eq.) were dissolved in 19.50 mL dry MeCN and cooled to 0 °C under Ar atmosphere. 146 μ L POCl₃ (1.57 mmol, 2.0 eq.) were added dropwise and the mixture was stirred at 0 °C for 10 min. Subsequently, 966 μ L NEt₃ (6.97 mmol, 8.9 eq.) were added and the mixture was stirred for another 20 min at 0 °C before 421 mg 2'-deoxyuridine derivative **8** (0.78 mmol, 1.0 eq.) were added. The reaction mixture was allowed to warm to room temperature and stirred for 18 h. After complete conversion the mixture was poured into 100 mL saturated aq. NaHCO₃ solution and extracted with CH₂Cl₂ (3 x 150 mL). Combined organic layers were dried over Na₂SO₄ and concentrated to dryness *in vacuo*.

The residue was dissolved in 15.60 mL 1,4-dioxane, 4.91 mL conc. NH₄OH were added and the mixture was stirred at 40 °C for 5 h. After cooling to room temperature the solution was poured into saturated aq. NH₄Cl solution (100 mL) and extracted with CH₂Cl₂ (3 x 100 mL). Combined organic layers were dried over Na₂SO₄, concentrated to dryness *in vacuo* and purified by silica gel column chromatography (CH₂Cl₂:MeOH 80:1 \rightarrow CH₂Cl₂:MeOH 60:1 \rightarrow CH₂Cl₂:MeOH 50:1 \rightarrow CH₂Cl₂:MeOH 30:1 \rightarrow CH₂Cl₂:MeOH 20:1 \rightarrow CH₂Cl₂:MeOH 15:1) to obtain 324 mg of the 2'-deoxycytidine derivative **9** (0.66 mmol, 84 %) as an off-white solid.

¹H-NMR (600 MHz, CDCl₃): δ /ppm = 9.96 (s, 1H, 2'''-H), 9.21 (s, 1H, 2'''-H), 7.91 – 7.84 (m, 4H, 3''-H), 7.24 – 7.16 (m, 4H, 4''-H), 6.67 (d, J = 9.6 Hz, 1H, 1'-H), 5.64 (dd, J = 6.5, 3.4 Hz, 1H, 3'-H), 4.66 – 4.48 (m, 6H, 4'-H, 5'-H, 1'''-H), 3.02 – 2.94 (m, 1H, 2'-H), 2.56 – 2.48 (m, 1H, 2'-H), 2.39 (s, 3H, 6''-H), 2.36 (s, 3H, 6''-H). **¹³C-NMR** (151 MHz, CDCl₃): δ /ppm = 166.45 (1''-C), 166.15 (1''-C), 154.50 (4-C), 144.47 (5''-C), 144.30 (5''-C), 137.91 (5-C), 129.94 (3''-C), 129.87 (3''-C), 129.42 (4''-C), 129.36 (4''-C), 126.91 (2''-C), 126.63 (2''-C), 87.22 (1'-C), 83.01 (4'-C), 75.07 (3'-C), 64.27 (5'-C), 60.98 (1'''-C), 35.38 (6''-C), 21.85 (6''-C). **HRMS (ESI⁺)**: calculated for C₂₅H₂₇N₄O₇⁺ [M+H]⁺ 495.1874, found: 495.1872; calculated for C₂₅H₂₆N₄O₇Na⁺ [M+Na]⁺ 517.1694, found: 517.1691. **HRMS (ESI⁻)**: calculated for C₂₅H₂₅N₄O₇⁻ [M-H]⁻ 493.1729; found: 493.1734. **IR (ATR)**: ν (cm⁻¹) = 3229 (br w), 1716 (s), 1611 (m), 1450 (w), 1272 (s), 1178 (m), 1104 (s), 1020 (m), 752 (s). **R_f** (CH₂Cl₂:MeOH 20:1): 0.26.

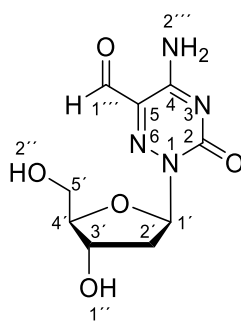
5-Formyl-(3',5'-di-O-p-toluoyl)-6-aza-2'-deoxycytidine **10**



In a heat dried Schlenk flask 306 mg of 5-hydroxymethyl-2'-deoxycytidine derivative **9** (0.62 mmol, 1.0 eq.) were dissolved in 18.6 mL dry CH_2Cl_2 and cooled to -15°C in a NaCl/ice cooling bath. Subsequently, 289 mg Dess-Martin periodinane (0.68 mmol, 1.1 eq.) were added, the mixture was allowed to warm to room temperature and stirred for 1 h at room temperature. After complete conversion the reaction was stopped by addition of a solution of 294 mg $\text{Na}_2\text{S}_2\text{O}_3$ (1.86 mmol, 3.0 eq.) in 100 mL saturated aq. NaHCO_3 and extracted with EtOAc (3 x 150 mL). Combined organic layers were dried over Na_2SO_4 , concentrated to dryness *in vacuo* and purified by silica gel column chromatography ($\text{CH}_2\text{Cl}_2 \rightarrow \text{CH}_2\text{Cl}_2:\text{MeOH}$ 40:1) to yield 272 mg of aldehyde **10** (0.55 mmol, 89 %) as a yellow foam.

$^1\text{H-NMR}$ (600 MHz, CDCl_3): δ/ppm = 9.40 (s, 1H, 1'''-H), 8.31 (s, 1H, 2'''-H), 7.94 – 7.86 (m, 4H, 3''-H), 7.28 – 7.20 (m, 4H, 4''-H), 6.87 (dd, J = 5.9 Hz, 1H, 1'-H), 5.74 (dd, J = 5.9, 2.9 Hz, 1H, 3'-H), 4.72 – 4.52 (m, 3H, 4'-H, 5'-H), 3.01 – 2.94 (m, 1H, 2'-H), 2.66 – 2.59 (m, 1H, 2'-H), 2.43 (s, 3H, 6''-H), 2.41 (s, 3H, 6'''-H). **$^{13}\text{C-NMR}$** (151 MHz, CDCl_3): δ/ppm = 188.81 (1'''-C), 166.26 (1''-C), 166.17 (1'-C), 155.17 (4-C), 144.69 (5''-C), 144.43 (5'-C), 129.97 (3''-C), 129.83 (3'-C), 129.44 (4''-C), 129.37 (5-C), 126.93 (2''-C), 126.48 (2'-C), 88.08 (1'-C), 83.34 (4'-C), 74.73 (3'-C), 63.87 (5'-C), 36.14 (2'-C), 21.89 (6''-C), 21.84 (6'''-C). **HRMS (ESI⁺)**: calculated for $\text{C}_{25}\text{H}_{25}\text{N}_4\text{O}_7^+$ $[\text{M}+\text{H}]^+$ 493.1718, found: 493.1717; calculated for $\text{C}_{25}\text{H}_{24}\text{N}_4\text{O}_7\text{Na}^+$ $[\text{M}+\text{Na}]^+$ 515.1537, found: 515.1534. **HRMS (ESI⁻)**: calculated for $\text{C}_{25}\text{H}_{23}\text{N}_4\text{O}_7^-$ $[\text{M}-\text{H}]^-$ 491.1572; found: 491.1577. **IR (ATR)**: ν (cm^{-1}) = 3237 (br w), 1720 (s), 1612 (m), 1272 (s), 1178 (w), 1104 (m), 753 (m). **R_f** ($\text{CH}_2\text{Cl}_2:\text{MeOH}$ 20:1): 0.34.

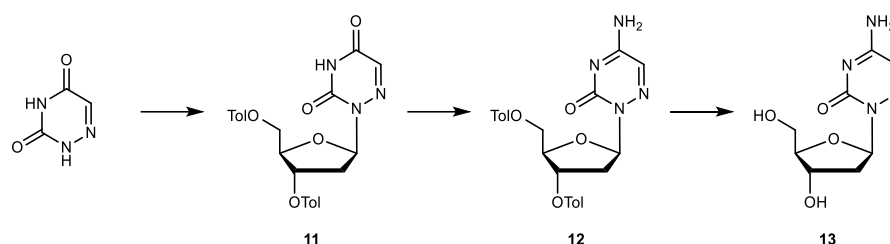
6-Aza-5-formyl-2'-deoxycytidine **3**



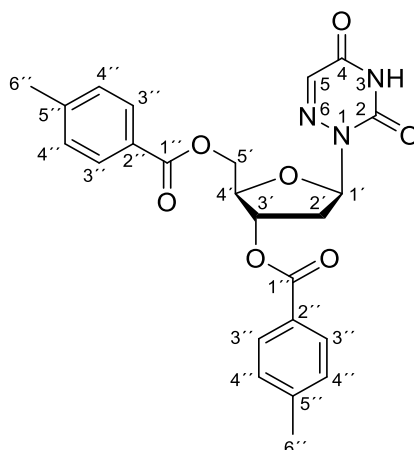
The deprotection was performed according to a modified procedure of *Mitchell et al.*^[2] 36 mg of toluoyl protected 6-Aza-nucleoside **10** (73 μ mol, 1.0 eq.) were dissolved in 0.73 mL dry MeOH and 0.73 dry benzene. To the solution 540 μ L 0.5 M NaOMe in dry MeOH (270 μ mol, 3.7 eq.) were added and the reaction mixture was stirred for 1.5 h at room temperature. After complete conversion the reaction mixture was neutralized with 2 M aq. HCl and evaporated to dryness. The resulting residue was dissolved in 10 mL ddH₂O and extracted with CH₂Cl₂ (5 x 10 mL). The aqueous layer was lyophilized and the residue was purified by *reversed phase* HPLC (0 % \rightarrow 3 % MeCN in H₂O in 45 min) to yield 10.1 mg of the desired product **3** (39 μ mol, 54 %) as a colorless solid.

¹H-NMR (600 MHz, DMSO-*d*₆): δ /ppm = 9.56 (s, 1H, 1'''-H), 8.61 (s, 1H, 2'''-H), 7.78 (s, 1H, 2'''-H), 6.52 (dd, J = 7.0, 5.0 Hz, 1H, 1'-H), 5.21 (d, J = 4.7 Hz, 1H, 1''-H), 4.63 (t, J = 5.9 Hz, 1H, 2''-H), 4.35 (ddd, J = 10.2, 5.3, 5.1 Hz, 1H, 3'-H), 3.76 (dd, J = 5.2, 5.1 Hz, 1H, 4'-H), 3.49 (dt, J = 11.2, 5.5 Hz, 1H, 5'-H), 3.41 (dt, J = 11.8, 6.0 Hz, 1H, 5'-H), 2.47 (ddd, J = 13.3, 6.3, 4.9 Hz, 1H, 2'-H), 2.15 (ddd, J = 13.1, 7.0, 5.3 Hz, 1H, 2'-H). **¹³C-NMR** (151 MHz, DMSO-*d*₆): δ /ppm = 189.77 (1'''-C), 156.56 (4-C), 151.66 (2-C), 129.61 (5-C), 87.87 (4'-C), 86.48 (1'-C), 70.50 (3'-C), 62.07 (5'-C), 37.74 (2'-C). **HRMS (ESI⁺)**: calculated for C₉H₁₃N₄O₅⁺ [M+H]⁺ 257.0881, found: 257.0880; calculated for C₉H₁₂N₄O₅Na⁺ [M+Na]⁺ 279.0700, found: 279.0698. **HRMS (ESI⁻)**: calculated for C₉H₁₁N₄O₅⁻ [M-H]⁻ 255.0735; found: 255.0735.

Synthesis of 6-Aza-2'-deoxycytidine



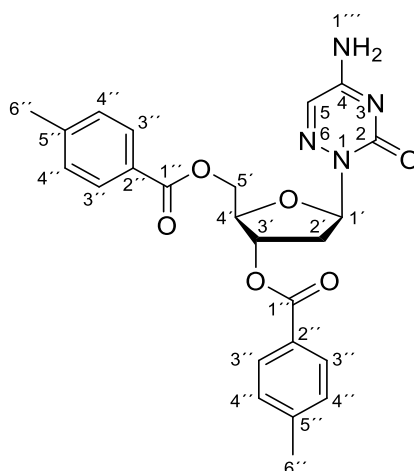
(3',5'-Di-O-p-toluoyl)-6-aza-2'-deoxyuridine 11



In a heat-dried round bottom flask and under argon atmosphere 300 mg 6-azauracil (2.65 mmol, 1.0 eq.) were suspended in 2.88 mL hexamethyldisilazane and 100 μ L TMSCl were added. The mixture was refluxed at 145 $^{\circ}$ C for 1 h and concentrated *in vacuo* at 30 $^{\circ}$ C after cooling to room temperature. The obtained dark oil was dried under high vacuum for 1 h and subsequently dissolved in 8.33 mL dry CHCl_3 . 1.13 g Hoffer's chlorosugar (2.92 mmol, 1.1 eq.) were added and the reaction mixture was stirred at room temperature for 16 h. The reaction mixture was poured into 250 mL saturated aq. NaHCO_3 solution and extracted with EtOAc (3 x 200 mL). Combined organic layers were dried over Na_2SO_4 and concentrated to dryness. Purification via silica gel column chromatography (iHex:EtOAc 3:1 \rightarrow iHex:EtOAc 1:1 \rightarrow iHex:EtOAc 1:2) yielded a mixture of diastereomeric β - and α - nucleosidation products (868 mg, 1.86 mmol, 70 %) as a colorless solid. Recrystallization from absolute EtOH yielded 742 mg of the pure β -nucleoside **11** (1.59 mmol, 60%) as a colorless solid.

$^1\text{H-NMR}$ (400 MHz, CDCl_3): δ /ppm = 8.82 (s, 1H, 3-H), 7.97 – 7.91 (m, 4H, 3''-H), 7.29 – 7.20 (m, 5H, 5-H, 4''-H), 6.68 (dd, J = 6.9, 6.7 Hz, 1H, 1'-H), 5.72 (ddd, J = 6.8, 3.9, 3.5 Hz, 1H, 3'-H), 4.66 (dd, J = 11.6, 4.3 Hz, 1H, 5'-H), 4.53 (dd, J = 9.7, 4.4 Hz, 1H, 4'-H), 4.46 (dd, J = 11.6, 5.0 Hz, 1H, 5'-H), 2.98 (dt, J = 13.4, 6.5 Hz, 1H, 2'-H), 2.48 (ddd, J = 14.1, 6.8, 3.9 Hz, 1H, 2'-H), 2.43 (s, 3H, 6''-H), 2.41 (s, 3H, 6''-H). **$^{13}\text{C-NMR}$** (101 MHz, CDCl_3): δ /ppm = 166.36 (1''-C), 166.16 (1''-C), 155.30 (4-C), 147.77 (2-C), 144.56 (5''-C), 144.11 (5''-C), 136.10 (5-C), 129.94 (3''-C), 129.90 (3''-C), 129.39 (4''-C), 129.27 (4''-C), 127.11 (2''-C), 126.61 (2''-C), 86.05 (1'-C), 82.73 (4'-C), 75.03 (3'-C), 64.06 (5'-C), 35.17 (2'-C), 21.90 (6''-C), 21.86 (6''-C). **HRMS (ESI $^+$)**: calculated for $\text{C}_{24}\text{H}_{24}\text{N}_3\text{O}_7^+$ $[\text{M}+\text{H}]^+$ 466.1609, found: 466.1610; calculated for $\text{C}_{24}\text{H}_{27}\text{N}_4\text{O}_7^+$ $[\text{M}+\text{NH}_4]^+$ 483.1874, found: 483.1874. **HRMS (ESI $^-$)**: calculated for $\text{C}_{24}\text{H}_{22}\text{N}_3\text{O}_7^-$ $[\text{M}-\text{H}]^-$ 464.1463; found: 464.1469. **IR (ATR)**: ν (cm^{-1}) = 3200 (w), 1727 (s), 1699 (s), 1610 (w), 1437 (w), 1392 (w), 1374 (w), 1329 (m), 1265 (s), 1173 (w), 1100 (s), 1080 (s), 954 (m), 808 (m), 752 (s). **R_f** (iHex: EtOAc 1:1): 0.44.

(3',5'-Di-O-p-toluoyl)-6-aza-2'-deoxycytidine 12

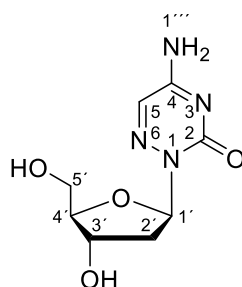


In a heat dried Schlenk-flask 668 mg 1,2,4-triazole (9.67 mmol, 9.0 eq.) were dissolved in 26.80 mL dry MeCN and cooled to 0 °C under Ar atmosphere. 201 μ L POCl₃ (2.15 mmol, 2.0 eq.) were added dropwise and the mixture was stirred at 0 °C for 10 min. Subsequently, 1.33 mL NEt₃ (9.56 mmol, 8.9 eq.) were added and the mixture was stirred for another 20 min at 0 °C before 500 mg of protected 6-aza-2'-deoxyuridine **11** (1.07 mmol, 1.0 eq.) were added. The reaction mixture was allowed to warm to ambient temperature and stirred for 17 h at room temperature. After complete conversion the mixture was poured into 250 mL saturated aq. NaHCO₃ solution and extracted with CH₂Cl₂ (3 x 200 mL). Combined organic layers were washed with brine (150 mL), dried over Na₂SO₄ and concentrated to dryness *in vacuo*.

The residue was dissolved in 21.40 mL 1,4-dioxane and 6.74 mL conc. NH₄OH were added. After stirring the mixture at 45 °C for 3 h the solution was poured into saturated aq. NH₄Cl solution (150 mL) and extracted with EtOAc (3 x 150 mL). Combined organic layers were washed with brine (150 mL), dried over Na₂SO₄ and concentrated to dryness *in vacuo*. The crude product was purified by silica gel column chromatography (CH₂Cl₂:MeOH 40:1 → CH₂Cl₂:MeOH 20:1 → CH₂Cl₂:MeOH 10:1) to obtain 460 mg of the amination product **12** (0.99 mmol, 93 %) as a slightly yellowish solid.

¹H-NMR (600 MHz, CDCl₃): δ /ppm = 8.57 (s, 1H, 1'''-H), 7.92 (m, 4H, 3''-H), 7.71 (s, 1H, 5-H), 7.32 (s, 1H, 1'''-H), 7.25 – 7.16 (m, 4H, 4''-H), 6.72 (dd, J = 6.6, 6.3 Hz, 1H, 1'-H), 5.73 (ddd, J = 7.2, 3.9, 3.1 Hz, 1H, 3'-H), 4.58 (ddd, J = 11.1, 5.4, 1.4 Hz, 1H, 5'-H), 4.55 – 4.52 (m, 1H, 4'-H), 4.47 (ddd, J = 11.2, 5.2, 1.2 Hz, 1H, 5'-H), 3.01 (ddd, J = 13.6, 6.8, 6.4 Hz, 1H, 2'-H), 2.45 (ddd, J = 13.9, 6.9, 4.2 Hz, 1H, 2'-H), 2.41 (s, 3H, 6''-H), 2.36 (s, 3H, 6'''-H). **¹³C-NMR** (151 MHz, CDCl₃): δ /ppm = 166.60 (1'-C), 166.20 (1''-C), 159.07 (4-C), 154.24 (2-C), 144.34 (5'''-C), 143.96 (5''-C), 129.94 (3''-C), 129.94 (3'''-C), 129.32 (4''-C), 129.21 (4'''-C), 128.61 (5-C), 127.10 (2''-C), 126.84 (2'''-C), 87.76 (1'-C), 82.23 (4'-C), 75.41 (3'-C), 64.55 (5'-C), 35.17 (2'-C), 21.85 (6''-C), 21.79 (6'''-C). **HRMS (ESI⁺)**: calculated for C₂₄H₂₅N₄O₆⁺ [M+H]⁺ 465.1769, found: 465.1770. **HRMS (ESI⁻)**: calculated for C₂₄H₂₃N₄O₆⁻ [M-H]⁻ 463.1623; found: 463.1630. **IR (ATR)**: ν (cm⁻¹) = 3305 (br w), 3059 (br w), 1715 (m), 1643 (m), 1610 (m), 1536 (w), 1466 (w), 1334 (w), 1265 (s), 1178 (m), 1102 (s) 1020 (m), 961 (w), 839 (w), 730 (s). **R_f** (CH₂Cl₂:MeOH 20:1): 0.30.

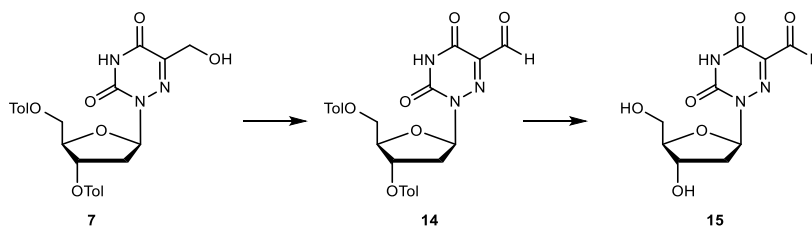
6-Aza-2'-deoxycytidine **13**



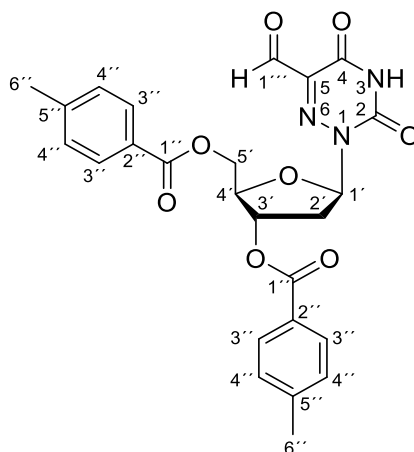
In a round bottom flask 10.0 mL of a solution of MeNH₂ in EtOH (33%) were added to 200 mg of toluoyl protected compound **12** (0.43 mmol, 1.0 eq.) and the mixture was stirred at room temperature for 14 h. As reaction control still showed protected starting material another 5.0 mL 33 % MeNH₂ in EtOH were added and the mixture was heated to 45 °C for 1 h. The solution was evaporated to dryness, dissolved in 10 mL ddH₂O and 10 mL CH₂Cl₂. Phases were separated, the aqueous layer was extracted with CH₂Cl₂ (4 x 10 mL) and subsequently concentrated *in vacuo*. The crude product was dissolved in ddH₂O (24 mL) and purified by reversed phase HPLC (0 % → 5 % MeCN in H₂O in 45 min) to yield 57.6 mg of the deprotected 6-Aza-nucleoside **13** as a white solid (0.25 mmol, 58 %).

¹H-NMR (600 MHz, D₂O): δ/ppm = 7.75 (s, 1H, 5-H), 6.57 (dd, *J* = 7.2, 4.8 Hz, 1H, 1'-H), 4.57 (ddd, *J* = 6.9, 4.4, 1.7 Hz, 1H, 3'-H), 4.05 – 4.01 (m, 1H, 4'-H), 3.85 – 3.69 (m, 1H, 5'-H), 3.71 – 3.61 (m, 1H, 5'-H), 2.72 (ddd, *J* = 13.9, 7.1, 4.8, 2.0 Hz, 1H, 2'-H), 2.35 (ddd, *J* = 14.7, 7.4, 4.0, 1.9 Hz, 1H, 2'-H). **¹³C-NMR** (151 MHz, D₂O): δ/ppm = 158.97 (4-C), 155.80 (2-C), 129.37 (5-C), 86.52 (1'-C), 86.44 (4'-C), 70.93 (3'-C), 61.80 (5'-C), 36.67 (2'-C). **HRMS (ESI)**: calculated for C₈H₁₁N₄O₄⁻ [M-H]⁻ 227.0786; found: 227.0787.

Synthesis of 6-Aza-5-formyl-2'-deoxyuridine



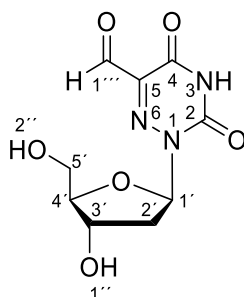
5-Formyl-(3',5'-di-O-p-toluoyl)-6-aza-2'-deoxyuridine **14**



In a Schlenk flask 158 mg of 5-hydroxymethyl-2'-deoxyuridine derivative **7** (0.32 mmol, 1.0 eq.) were dissolved in 9.6 mL dry CH_2Cl_2 and cooled to -15°C . At this temperature 162 mg Dess-Martin periodinane (0.38 mmol, 1.2 eq.) were added, the mixture was allowed to warm to room temperature slowly and was further stirred for 2.5 h. The reaction was stopped by addition of a solution of 182 mg $\text{Na}_2\text{S}_2\text{O}_3$ (1.15 mmol, 3.6 eq.) in 50 mL saturated aq. NaHCO_3 and the mixture was extracted with EtOAc (3 x 100 mL). Combined organic layers were dried over Na_2SO_4 , concentrated to dryness *in vacuo* and purified by silica gel column chromatography (iHex:EtOAc 1:1 \rightarrow EtOAc) to yield 152 mg of aldehyde **14** (0.31 mmol, 96 %) as an off-white foam.

$^1\text{H-NMR}$ (400 MHz, CDCl_3): δ/ppm = 9.58 (s, 1H, 1'''-H), 9.02 (s, 1H, 3H), 8.01 – 7.79 (m, 4H, 3'-H), 7.29 – 7.18 (m, 4H, 4''-H), 6.70 (t, J = 6.3 Hz, 1H, 1'-H), 5.72 (dd, J = 6.7, 3.4 Hz, 1H, 3'-H), 4.70 – 4.49 (m, 3H, 4'-H, 5'-H), 3.07 (dt, J = 14.2, 6.4 Hz, 1H, 2'-H), 2.58 (ddd, J = 14.2, 6.7, 4.0 Hz, 1H, 2'-H), 2.43 (s, 3H, 6''-H), 2.40 (s, 3H, 6''-H). **$^{13}\text{C-NMR}$** (101 MHz, CDCl_3): δ/ppm = 184.16 (1'''-C), 166.33 (1''-C), 166.14 (1''-C), 152.79 (4-C), 147.15 (2-C), 144.68 (5''-C), 144.40 (5''-C), 137.22 (5-C), 129.94 (3''-C), 129.81 (3''-C), 129.42 (4''-C), 129.40 (4''-C), 126.81 (2''-C), 126.41 (2''-C), 87.16 (1'-C), 83.20 (4'-C), 74.59 (3'-C), 63.74 (5'-C), 35.40 (2'-C), 21.91 (6''-C), 21.85 (6''-C). **HRMS (ESI $^+$)**: calculated for $\text{C}_{25}\text{H}_{27}\text{N}_4\text{O}_8^+$ $[\text{M}+\text{NH}_4]^+$ 511.1823, found: 511.1823. **HRMS (ESI $^-$)**: calculated for $\text{C}_{25}\text{H}_{22}\text{N}_3\text{O}_8^-$ $[\text{M}-\text{H}]^-$ 492.1412; found: 492.1413. **IR (ATR)**: ν (cm^{-1}) = 2963 (w), 1712 (s), 1611 (m), 1439 (w), 1398 (w), 1309 (w), 1260 (s), 1178 (m), 10963 (s), 1020 (s), 909 (m), 800 (s), 752 (s), 732 (s). **R_f** (iHex: EtOAc 1:1): 0.17.

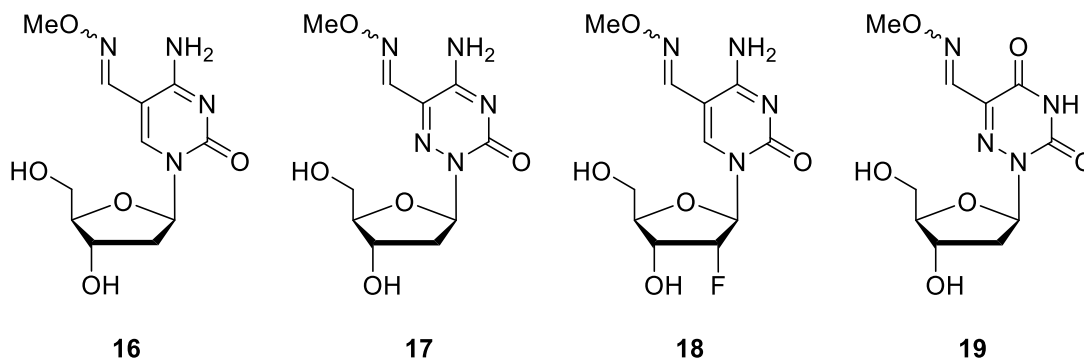
5-Formyl-6-aza-2'-deoxyuridine **15**



97 mg of toluoyl protected 6-Aza-nucleoside **14** (0.20 mmol, 1.0 eq.) were dissolved in 2.0 mL dry MeOH and 2.0 dry benzene. To the solution freshly prepared 0.5 M NaOMe in dry MeOH (1.46 mL, 0.73 mmol, 3.7 eq.) was added and the reaction mixture was stirred for 2 h at room temperature. Subsequently, the reaction mixture was neutralized with 2 M aq. HCl, evaporated to dryness and redissolved in 10 mL ddH₂O. The mixture was extracted with CH₂Cl₂ (4 x 10 mL). The aqueous layer was lyophilized and the residue was purified by *reversed phase* HPLC (0 % → 3 % MeCN in H₂O in 45 min) to yield 16.3 mg of the product **15** (63 μmol, 32 %) as a colorless solid.

¹H-NMR (400 MHz, D₂O): δ/ppm = 9.62 (s, 0.2H, 1'''-H), 6.55 (dd, *J* = 7.2, 4.5 Hz, 0.2H, 1'-H), 6.48 (dd, *J* = 7.2, 5.2 Hz, 0.8H, 1''-H), 5.87 (s, 0.7H, 1'''-H), 4.62 – 4.48 (m, 1H, 3'-H), 3.96 (m, 1H, 4'-H), 3.75 – 3.55 (m, 2H, 5'-H), 2.71 (m, 1H, 2'-H), 2.28 (m, 1H, 2'-H). **¹³C-NMR** (101 MHz, D₂O): δ/ppm = 190.31 (1'''-C), 165.12 (4-C), 162.87 (4-C), 156.31 (2-C), 155.12 (2-C), 142.78, 137.09 (5-C), 87.81 (4'-C), 86.93 (4'-C), 86.48 (1'-C), 86.42 (1'-C), 86.14 (1'''-C), 71.05 (3'-C), 70.83 (3'-C), 61.81 (5'-C), 61.70 (5'-C), 36.89 (2'-C), 36.57 (2'-C). **HRMS (ESI)**: calculated for C₉H₁₀N₃O₆⁻ [M-H]⁻ 256.0575; found: 256.0576.

Reactions of **1**, **2**, **3** and **15** with Methoxyamine



To a solution of 100 nmol nucleoside (1.0 eq.) in 35.0 μL aq. NaOH (pH = 10) was added 33.33 μL aq. MeONH₂ (150 mM, 50.0 eq.) and the mixture was incubated at 25 °C for 1 h. The reaction was stopped by the addition of 35.0 μL of aq. HCOOH (pH = 3) and analyzed by reversed phase HPLC. The resulting reaction products were collected and analyzed via HRMS.

	16	17	18	19
Conversion of SM	50 %	quant.	57 %	quant.
HPLC gradient [MeCN in H ₂ O in 45 min]	0 % \rightarrow 15 %	0 % \rightarrow 30 %	0 % \rightarrow 13 %	0 % \rightarrow 30 %
HRMS (calcd.) ESI ⁽⁺⁾	285.1194 C ₁₁ H ₁₇ O ₅ N ₄ ⁺ [M+H] ⁺	286.1146 C ₁₀ H ₁₆ O ₅ N ₅ ⁺ [M+H] ⁺	303.1099 C ₁₁ H ₁₆ O ₅ N ₄ F ⁺ [M+H] ⁺	
ESI ⁽⁻⁾	283.1048 C ₁₁ H ₁₅ O ₅ N ₄ ⁻ [M-H] ⁻	284.1000 C ₁₀ H ₁₄ O ₅ N ₅ ⁻ [M-H] ⁻		285.0841 C ₁₀ H ₁₃ O ₆ N ₄ ⁻ [M-H] ⁻
HRMS (found) ESI ⁽⁺⁾	285.1193 [M+H] ⁺	286.1151 [M+H] ⁺	303.1098 [M+H] ⁺	
ESI ⁽⁻⁾	283.1048 [M-H] ⁻	284.1000 [M-H] ⁻		285.0840 [M-H] ⁻

Reaction of **1**, **2** and **3** with NaHSO₃

20 mM nucleoside solution in ddH₂O was shaken with 4.36 M aq. NaHSO₃ (435 eq, pH = 5) at 55 °C. Samples were taken at time points of 0, 5 min, 20 min, 30 min, 40 min, 1 h, 2 h, 3 h, 4 h, 5 h, 6 h and 15 h. Bisulfite was cleaved by preparing 1 mM solution of reaction mixture and 2 M NaOH (1700 eq), shaken for 10 min at rt and followed by immediate injection to HPLC (Gradients: **1** 0% to 13% MeCN in H₂O in 45 min, **2** 0% to 15% MeCN in H₂O in 45 min, **3** 0% to 3% MeCN in H₂O in 45 min).

Chromatograms of each time point were normalized to the integral of the reagent and corrected accordingly. Conversion was determined by the decrease of the corrected integrals of the starting materials.

Cell culture

DMEM high glucose (Sigma Aldrich) containing 10% heat-inactivated FBS (Gibco), 100 U/mL penicillin, 100 µg/mL streptomycin, 2 mM L-glutamine were used to culture Neuro-2a and RBL-2H3 cell lines. DMEM high glucose and Ham's Nutrient Mixture F12 (Sigma Aldrich) containing 10% heat-inactivated FBS (Gibco), 100 U/mL penicillin, 100 µg/mL streptomycin, 2 mM L-glutamine were used to culture CHO-K1 cells. Metabolic-labelling experiments were conducted by plating cells in their dedicated medium containing 350 µM of F-fdC or a-fdC (in case of the co-feeding study: 350 µM of each) for three days.

DMEM high glucose (Sigma Aldrich) containing 10% FBS (PAN Biotech), 2 mM L-glutamine, 1x MEM Non-essential Amino Acid Solution and 0.1 mM β-mercaptoethanol (Sigma Aldrich) were used as basal medium for E14 TDG +/- and -/- (obtained from Cortázar *et al.*) mESC cultures. The mESC lines were maintained in naïve state on gelatin coated plates by supplementing basal medium with 1000 U/mL LIF (ORF Genetics), 3.0 µM GSK3 inhibitor CHIR99021 and 1.0 µM MNK inhibitor CGP57380 (a2i medium). Metabolic labelling experiments with isotope-labeled nucleosides were performed by plating mESCs in priming conditions consisted of basal mESC medium supplemented with 3. µM GSK3 inhibitor CHIR99021 and 2.5 µM tankyrase inhibitor IWR1-endo. Labelled nucleosides were added at the concentration of 350 µM each to the priming medium over three days.

All samples were washed with PBS (Sigma Aldrich) once before harvesting and lysed directly in the plates by adding RLT buffer (Qiagen) supplemented with 400 µM 2,6-di-tert-butyl-4-methylphenol (BHT) and desferoxamine mesylate (DM). Next, DNA was sheared in MM400 bead mill (Retsch) at 30 Hz for 1 min in 2 mL microcentrifuge tubes (Eppendorf) with 5 mm diameter stainless steel beads (one per tube) and centrifuged at 21000rcf for 5 minutes. Genomic DNA (gDNA) was extracted using Zymo Quick gDNA mini-prep® kit according to the protocol with an addition of RNase A treatment (Qiagen) at 0.2 mg/mL in Genomic Lysis Buffer – 400 µL of the solution transferred directly on the column and incubated for 15 minutes. All samples were eluted in DNase-free ddH₂O with 20 µM BHT, the concentration of gDNA was measured on Nanodrop.

While investigating the deformylation levels using F-fdC as a metabolic label, the following conditions were used for the gDNA digestion procedure: Degradase (1.5 µL) per up to 10 µg of F-fdC labelled gDNA in 35 µL in H₂O. The digestion mixture was incubated at 37 °C for 4 hours. Then, the samples were filtered using AcroPrep Advance 0.2 µm 96-well filter plate (Pall Life Sciences) prior to LC-MS/MS analysis (39 µL injection volume at 4 °C). This method was compared to NEB enzyme mix digestion (described below) and showed no

difference in the QQQ measurement outcome, thus we concluded that both can be used at equal efficiency.

Chemical labelling and spiking

Chemical labelling of a-fdC was needed to receive a sharper signal during the MS analysis. Therefore, methoxyamine was used as a derivatization reagent. We optimized the reaction conditions including time and temperature.

The quantification of the the Fluoro compounds as well as the other known nucleosides was done by spiking of heavy labelled nucleosides which were synthesized by ourselves.

For the determination of a-fdC in the gDNA samples we did an external calibration.

Analysis of labelled and unlabelled products of F-fdC and a-fdC

The analysis was performed using an UHPLC-QQQ-MS/MS system consisting of a Triple Quad™ 6490 mass spectrometer (Agilent) with an ESI source and an Agilent Infinity 1290 UHPLC. The elution was monitored at 260 nm (Agilent InfinityLab Deuterium Lamp G1314). Data Acquisition and processing were performed using MassHunter Workstation Software Version B.07.01 (Agilent).

The UHPLC separation was performed for the Fluoro-fdC as well as mdC, hmdC, fdC on an InfinityLab Poroshell 120 SB-C8 column (2.1 mm x 150 mm, 2.7 µm, Agilent Technologies, USA) at 35 °C. Water containing 0.0085% FA (v/v, solvent A) and MeCN containing 0.0085% FA (v/v, solvent B) was used as the mobile phase. A gradient of 0 - 3.5% B for 4 min, 3.5 - 5% B for 2.9 min, 5 - 80% B for 0.3 min, 80% B for 3.3 min was used. The flow rate of the mobile phase was set to 0.35 mL min⁻¹.

The derivatized a-fdC, F-fdC and fdC were separated on an InfinityLab Poroshell 120 SB-C18 column (2.1 mm x 150 mm, 2.7 µm, Agilent Technologies, USA) at 35 °C. Water containing 0.0085% FA (v/v, solvent A) and MeCN containing 0.0085% FA (v/v, solvent B) was used as the mobile phase. A gradient of 0% B for 3 min, 0 - 15% B for 27 min, 15 - 100% B for 5 min was used. The flow rate of the mobile phase was set to 0.35 mL min⁻¹.

The mass spectrometry detection was performed under positive ESI mode. The nucleosides and labelled products were monitored using the multiple reaction monitoring (MRM) mode. The MRM parameters were optimized to achieve maximal detection sensitivity (Tables 2 and 3).

Table 2 MRM parameters for the detection of F-fdC and its derivatives

Name	Transition	Scan	Type	Precursor Ion	Product Ion	Ion Polarity
UV-dG	0,0 -> 0,0	MRM	Target	0,0	0,0	Positive
UV-dC	0,0 -> 0,0	MRM	Target	0,0	0,0	Positive
UV total	0,0 -> 0,0	MRM	Target	0,0	0,0	Positive
Fluoro-fdC-dN	274,1 -> 140,1	MRM	Target	274,1	140,1	Positive
Fluoro-fdC-dN- ¹⁵ N ₂	276,1 -> 142,0	MRM	ISTD	276,1	142,0	Positive
Fluoro-dU	245,1 -> 225,1	MRM	Target	245,1	225,1	Negative
Fluoro-dC	246,1 -> 112,1	MRM	Target	246,1	112,1	Positive
F-fdU	273,1 -> 253,1	MRM	Target	273,1	253,1	Negative
fdC-dN- ¹⁵ N ₂	258,1 -> 142,0	MRM	ISTD	258,1	142,0	Positive
fdC-dN	256,1 -> 140,1	MRM	Target	256,1	140,1	Positive
¹⁵ N ₂ -FdC	248,1 -> 114,0	MRM	ISTD	248,1	114,0	Positive

Table 3 MRM parameters for the derivatized a-fdC samples

Name	Transition	Scan	Type	Precursor Ion	Product Ion	Ion Polarity
MeON-a-fdC	286,1 -> 170,1	MRM	Target	286,1	170,1	Positive
a-fdC	257,1 -> 141,1	MRM	Target	257,1	141,1	Positive
a-dC	229,1 -> 188,1	MRM	Target	229,1	188,1	Positive
a-dC	229,1 -> 112,1	MRM	Target	229,1	112,1	Positive

Digestion of the DNA of N2a cell line

As shown in Table 4 we used the following chemicals for the analysis of our samples. Reaction buffer 10X and Enzyme mix was bought as a Nucleoside Digestion Mix (M0649S) kit (*New England BioLabs Inc.*). The nucleosides mix was prepared by ourselves and contains heavy labelled mdC, hmdC, fdC, cadC, 8oxodG and hmdU. Furthermore we spiked heavy labelled Fluoro-dC and Fluoro-fdC for later quantification of the Fluoro-compounds. The concentration of the DNA for the digestion was received from the Nanodrop. The total volume was then incubated at 37 °C for 1.5 hours.

Table 4: Digestion sheet of the gDNA samples

	Sample	c(ng/ μ L)	m (DNA) [ng]	V (H ₂ O) [μ L]	V (DNA in H ₂ O) [μ L]	V (Reaction buffer 10X)	V (Enzyme mix)	V (Nucleosides mix)	Spiking heavy Fluoro-dC	Spiking heavy Fluoro-fdC	Total Volume [μ L]	1 h 30 min 37 °C	Injection volume [μ L]
Blank 1	1	0	0	38,4	0,0	5,0	3,0	3,0	0,35	0,25	50,0		39
Blank 2	2	0	0	38,4	0,0	5,0	3,0	3,0	0,35	0,25	50,0		39
Control N2a Cofeeding 10 μ g	3	528	10000	19,5	18,94	5,0	3,0	3,0	0,35	0,25	50,0		39
Control N2a Cofeeding 10 μ g	4	528	10000	19,5	18,94	5,0	3,0	3,0	0,35	0,25	50,0		39
Control N2a Cofeeding 5 μ g	5	528	5000	28,9	9,47	5,0	3,0	3,0	0,35	0,25	50,0		39
Control N2a Cofeeding 5 μ g	6	528	5000	28,9	9,47	5,0	3,0	3,0	0,35	0,25	50,0		39
Sample N2a Cofeeding 10 μ g	7	412	10000	14,1	24,27	5,0	3,0	3,0	0,35	0,25	50,0		39
Sample N2a Cofeeding 10 μ g	8	412	10000	14,1	24,27	5,0	3,0	3,0	0,35	0,25	50,0		39
Sample N2a Cofeeding 5 μ g	9	412	5000	26,3	12,14	5,0	3,0	3,0	0,35	0,25	50,0		39
Sample N2a Cofeeding 5 μ g	10	412	5000	26,3	12,14	5,0	3,0	3,0	0,35	0,25	50,0		39

Derivatization of a-fdC in digested gDNA of N2a cell line

Derivatization of a-fdC from the digested DNA (5 µg) of biological samples by methoxyamine was performed under the following conditions. Briefly, the nucleoside mixture was derivatized in 50 µL H₂O with 40 µL aq. NaOH (pH 10) and 8 µL methoxyamine (163 mM in H₂O) for 45 min at 25 °C. Afterwards the solution was neutralized with 40 µL aq. FA (pH 3). The derivatized nucleosides were then lyophilized and resuspended in 50 µL H₂O.

After resuspension the samples were filtered utilizing a 0.2 µm Supor filtration plate (*Pall Corporation*) and subjected to UHPLC-QQQ-MS/MS.

The determination of F-fdC and some other nucleosides in gDNA (10 µg) was performed by using heavy labelled compounds. After the digestion the samples were filtered utilizing a 0.2 µm Supor filtration plate (*Pall Corporation*) and subjected to UHPLC-QQQ-MS/MS.

Calibration curve for external quantification of a-fdC in gDNA of N2a cell lines

The quantification of a-fdC was done by an external calibration.^[3] The external calibration curve was done by serially diluting pure a-fdC (see Table 5) and measured in technical triplicates prior to each measurement. Linear regression was done by Microsoft Excel (Figure 6). The injection volume was 29 µL.

Table 5 Measured values for the external calibration curve of a-fdC

n [fmol]	Average area	standard deviation
0,11	1006	110
0,22	1822	98
0,44	3160	414
0,88	4754	428
1,77	8439	333
3,53	14692	346
7,07	28982	522
14,14	56445	1107
28,28	111343	2041
56,55	220616	2561
113,10	439962	2460

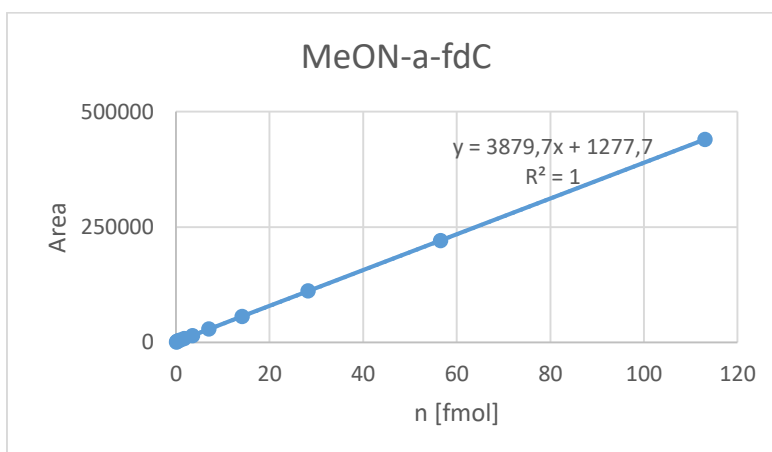


Figure 6: Calibration curve for the external quantification of a-fdC

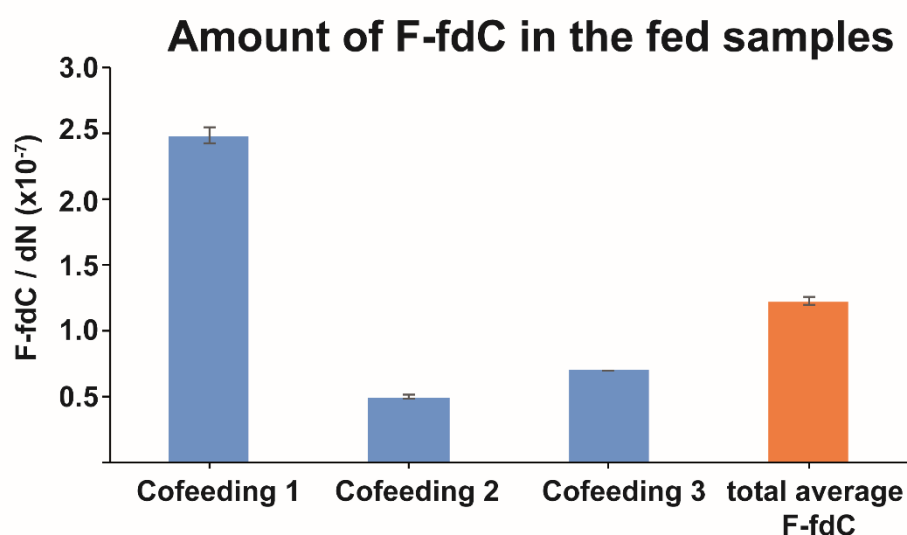
Quantification of concentration of the fed compounds

The quantification of the fed compounds was performed in two technical replicates of a biological triplicate.

For the quantification of a-fdC one has to say, that there was a small background in the a-fdC peak, which was continuously subtracted from the a-fdC value. This value was determined by the control sample.

The quantification of the Fluoro-fdC and the deformylated Fluoro-dC was done over the ratio of unlabelled to labelled compound. Furthermore it was compared to the amount of deoxycytidine and deoxy-Guanosine and dN. The final results can be seen in Figures 7.

A



B

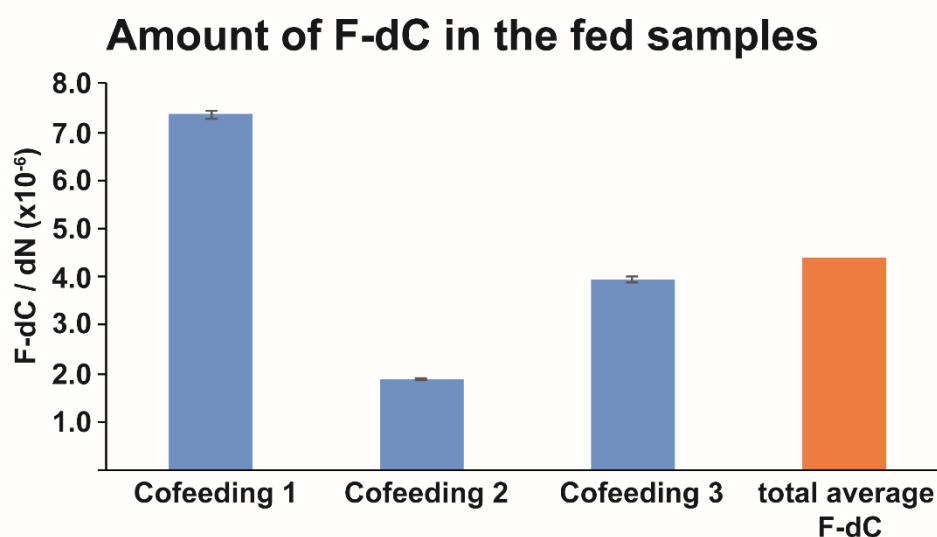


Figure 7 Measured amount of F-fdC (A) and F-dC (B) in the fed samples of N2a cells

Furthermore we were able to calculate the amount of deformylation of Fluoro-fdC by the amounts of Fluoro-fdC and Fluoro-dC. This was able to be done by the unnatural Fluoro compound. In Figure 8 the deformylation rates are shown.

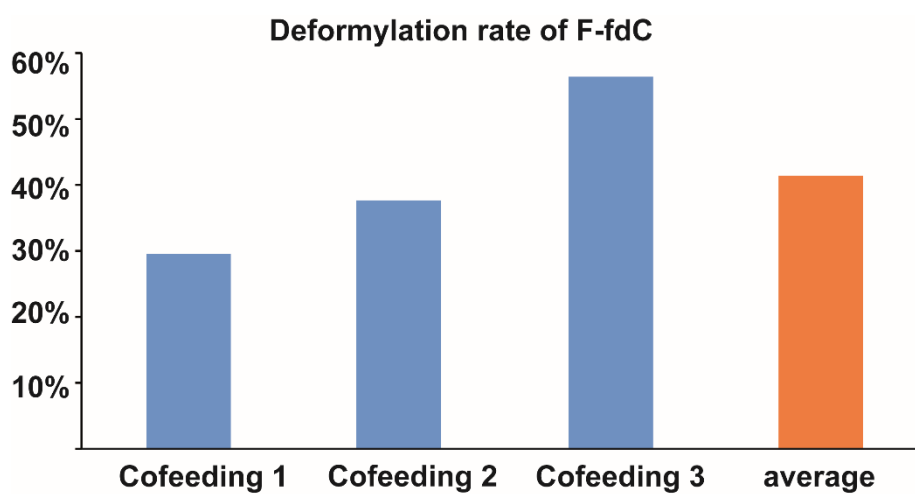
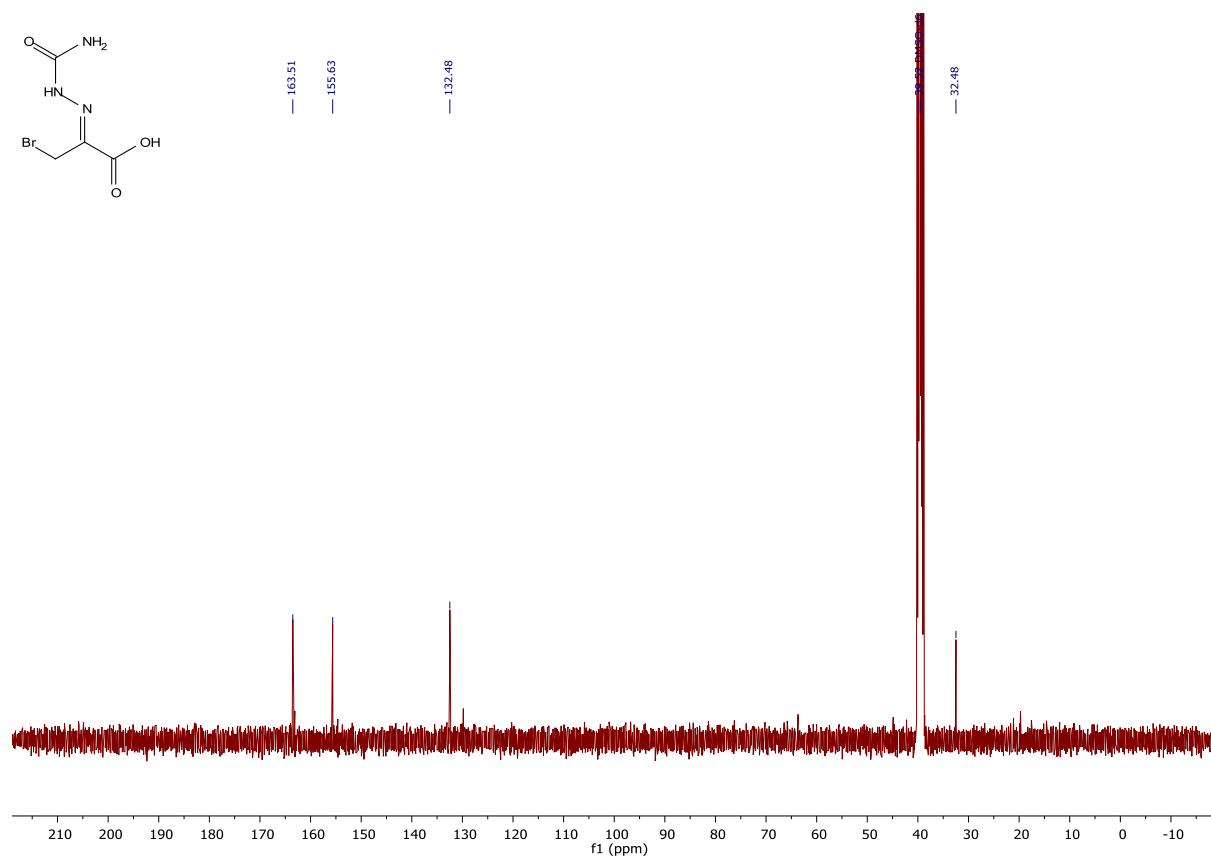
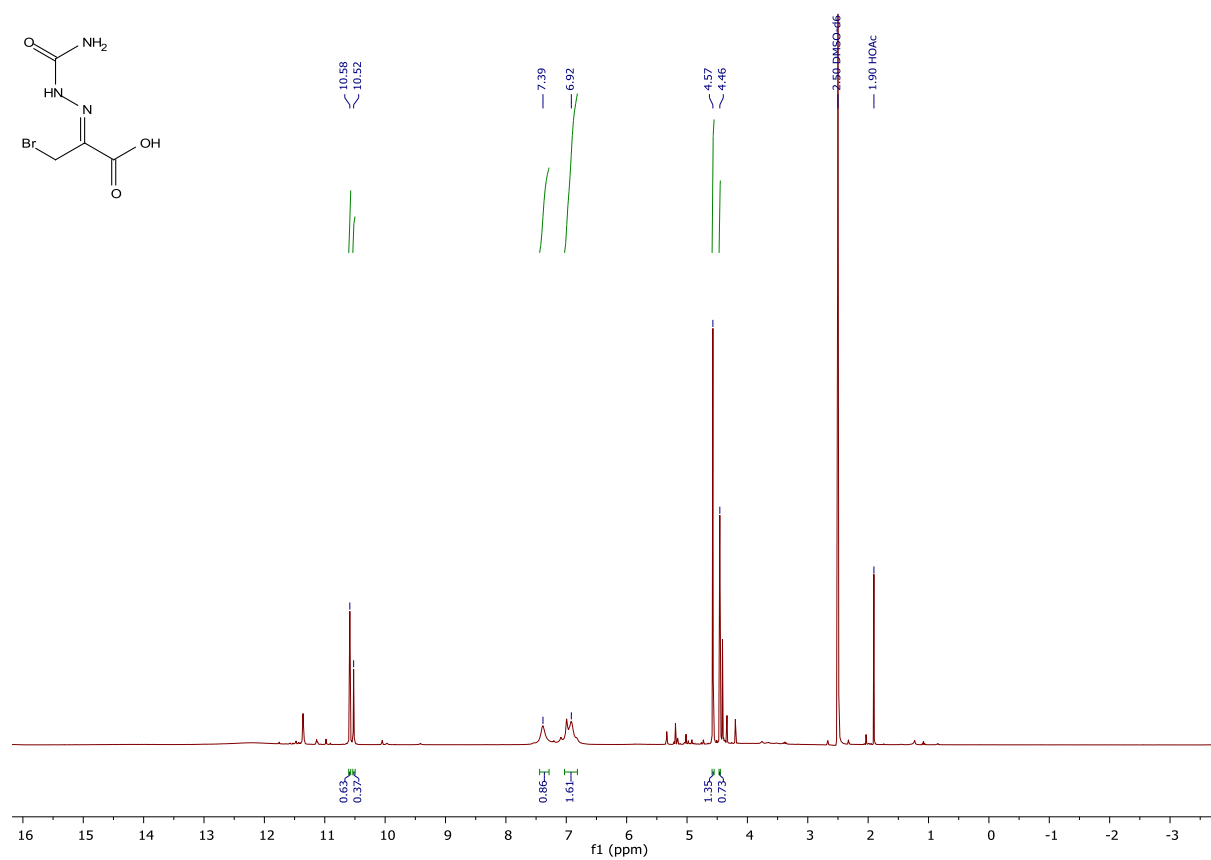
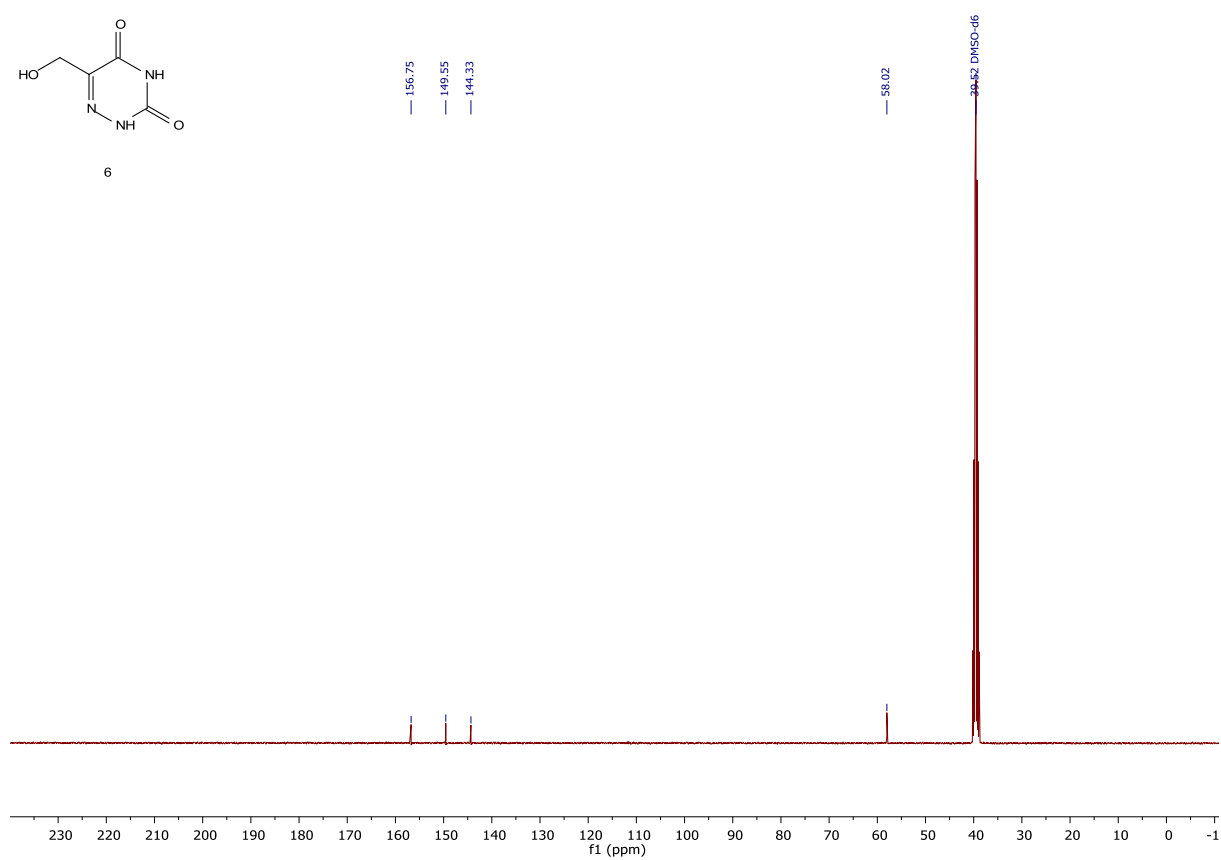
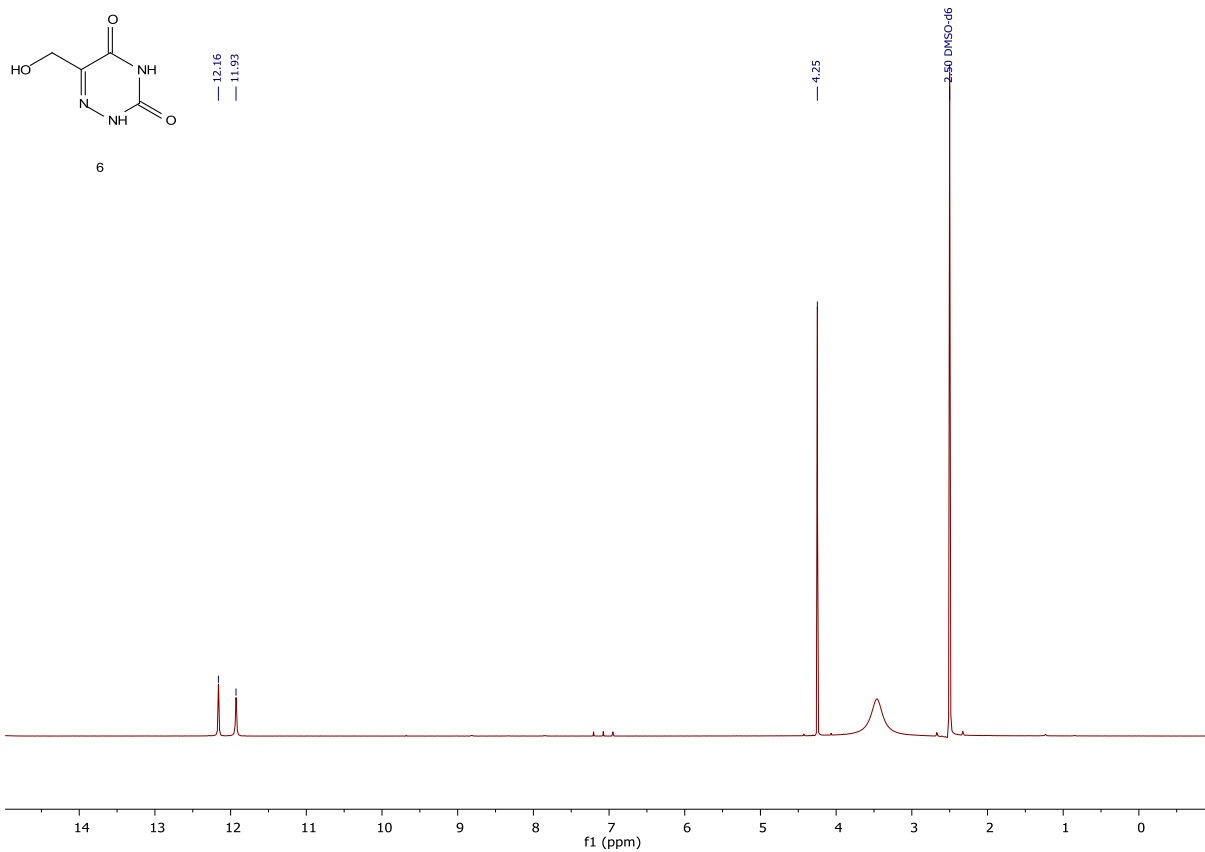
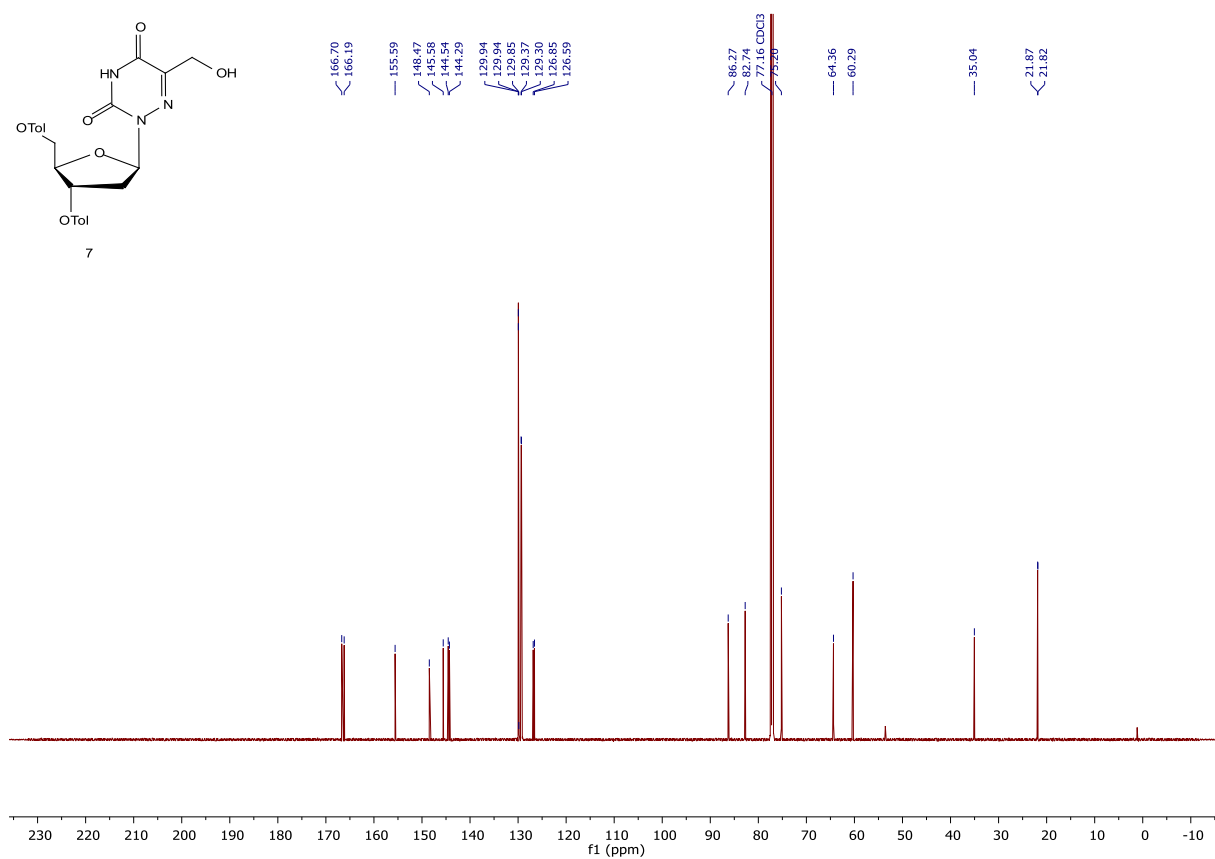
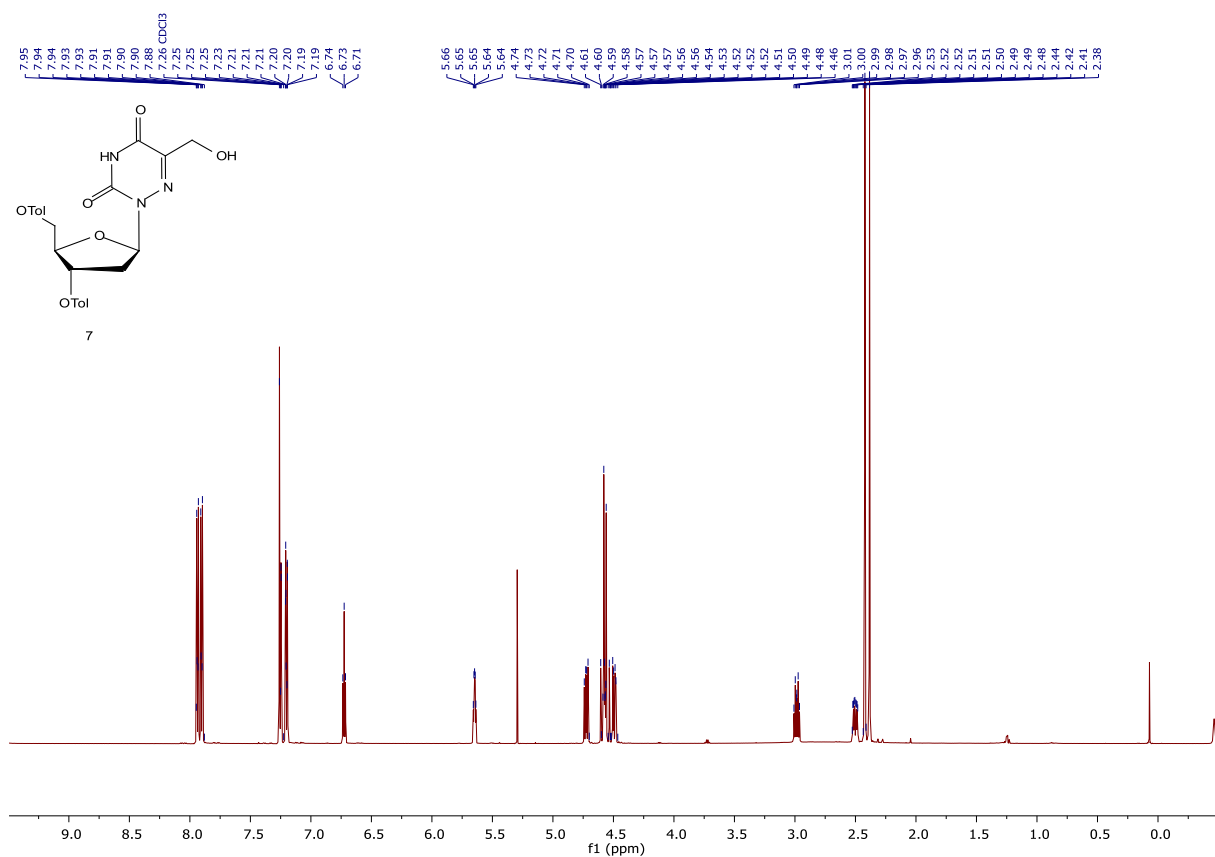


Figure 8 Calculated deformylation rate of F-fdC in N2a cells

NMR spectra of synthesized compounds

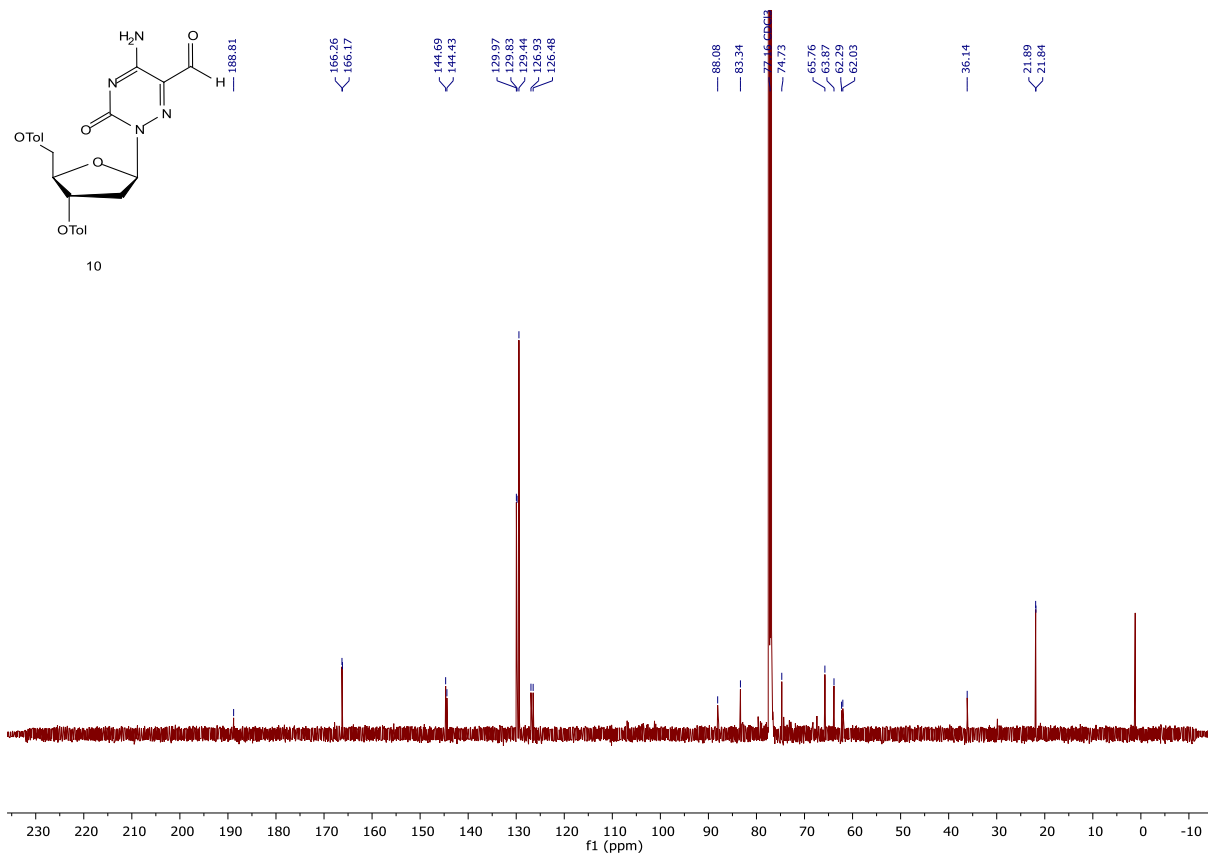
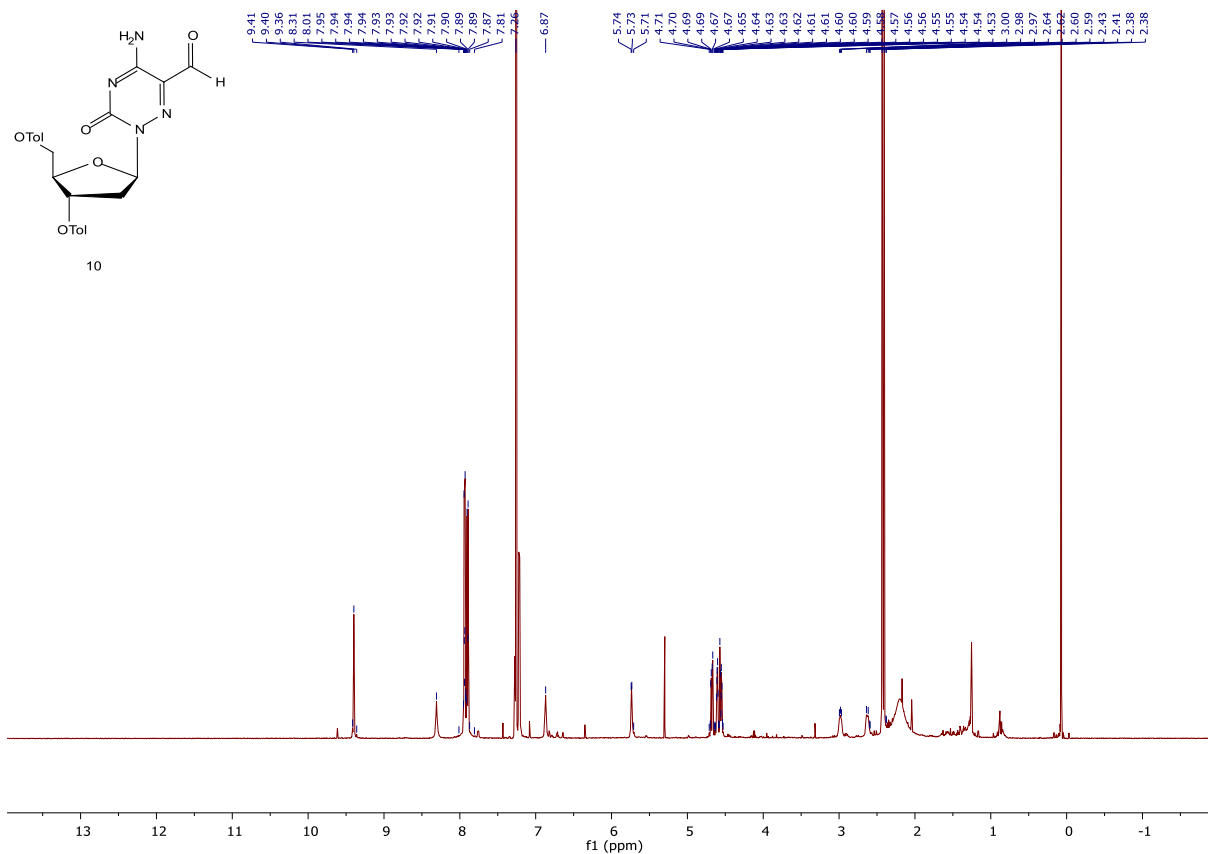


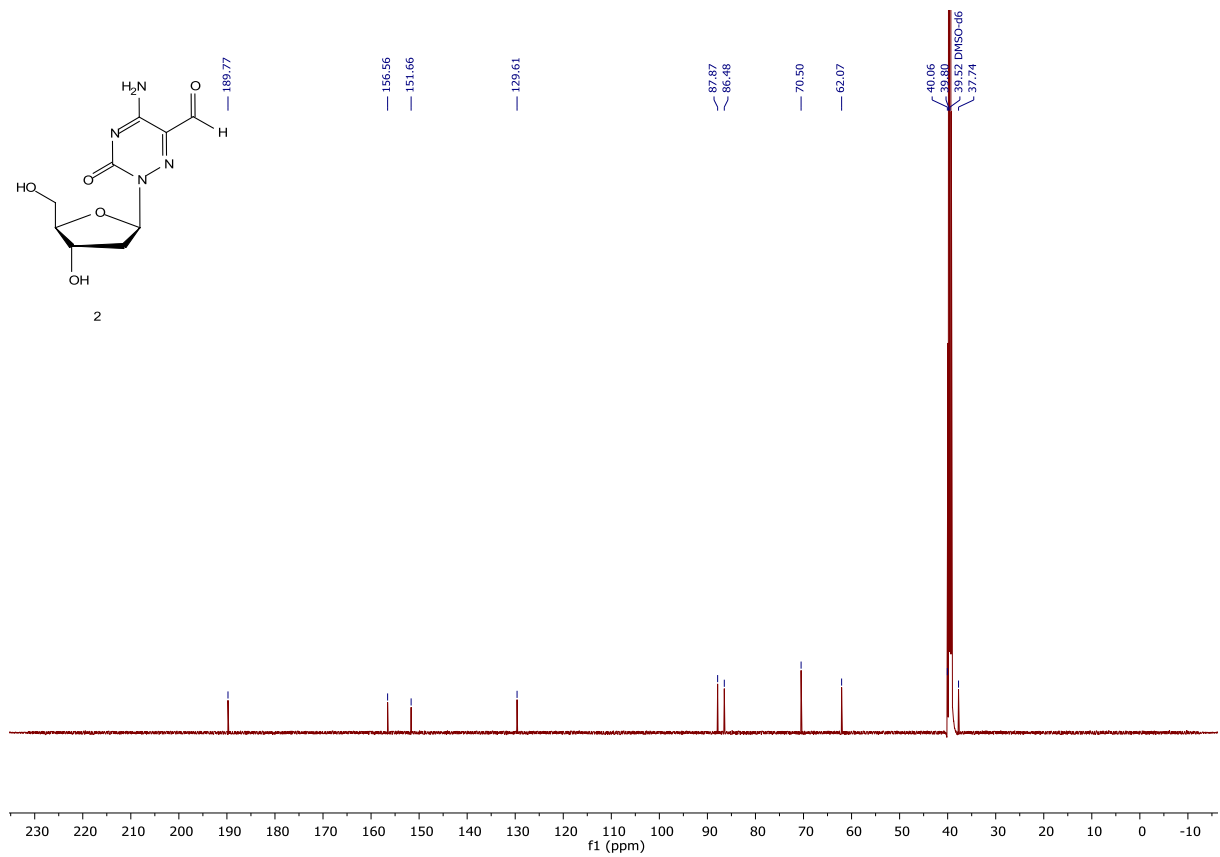
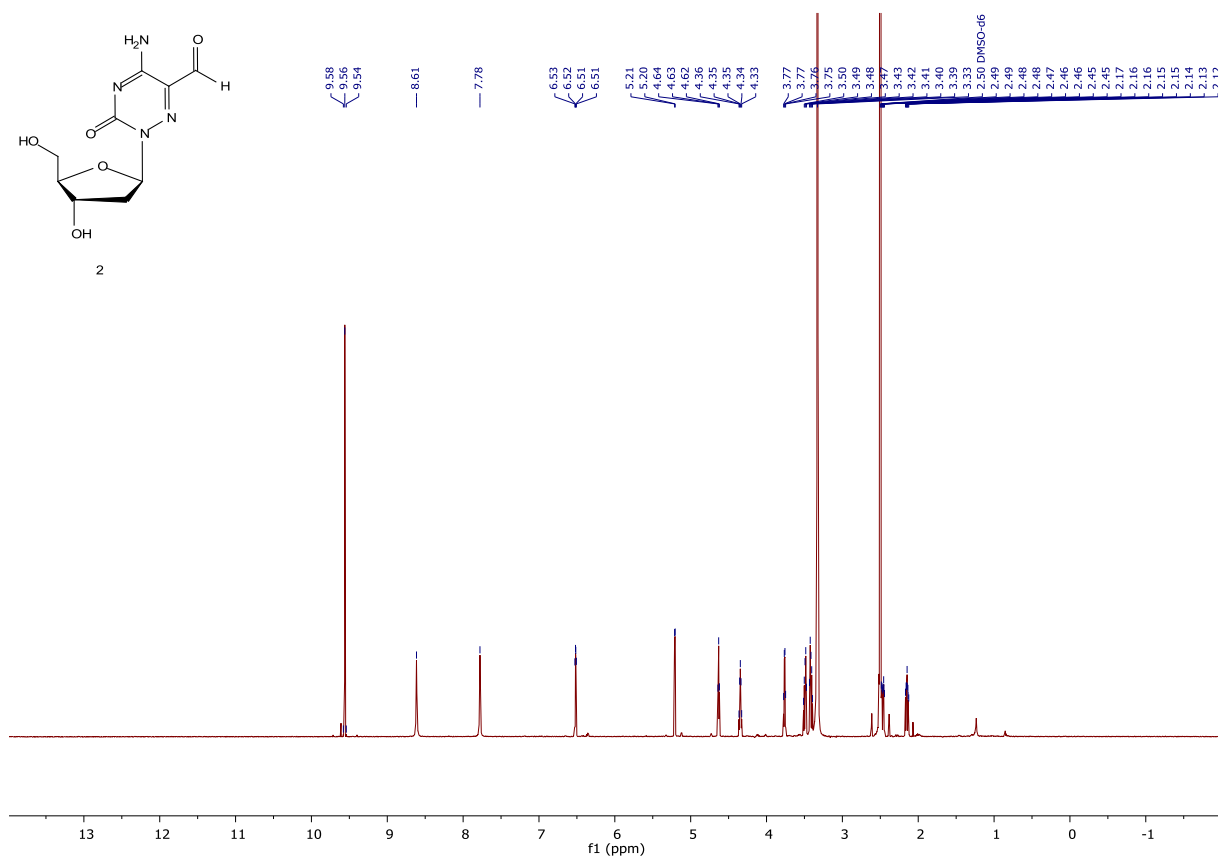


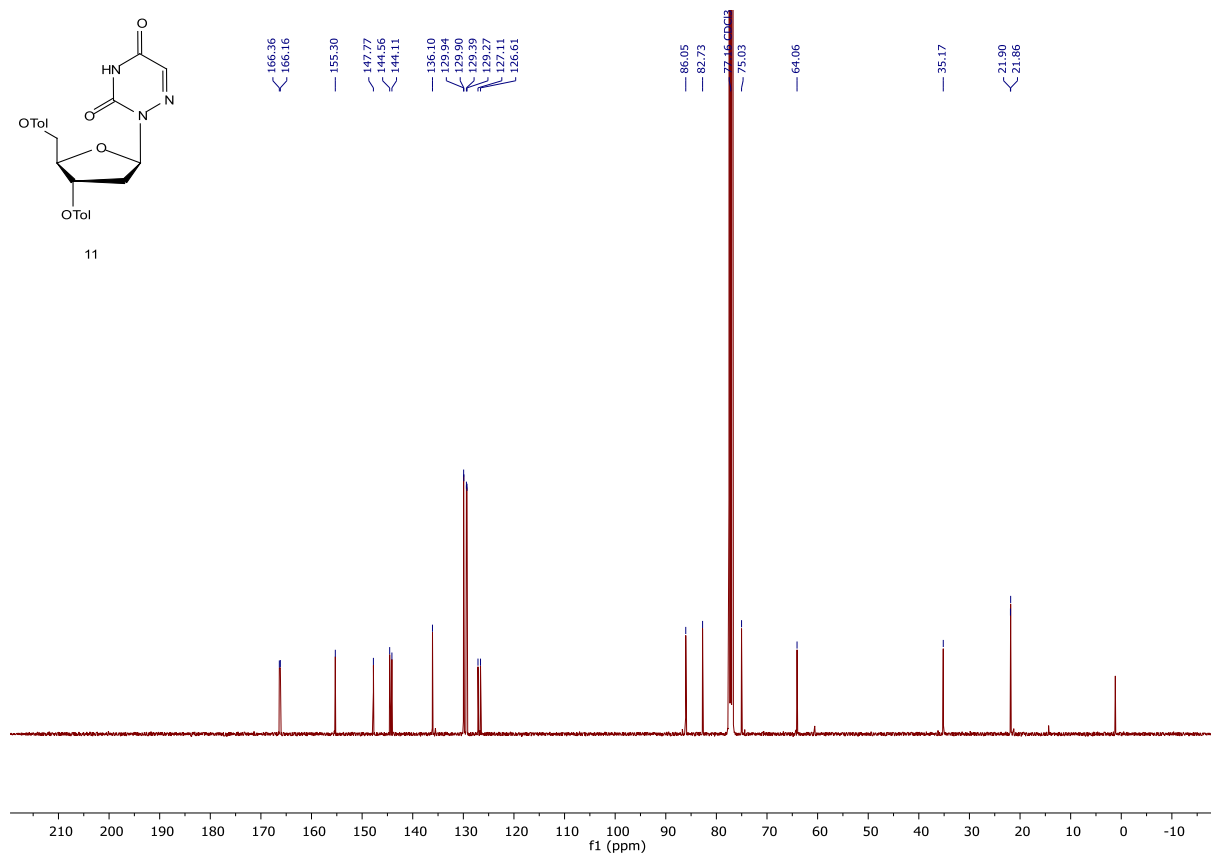
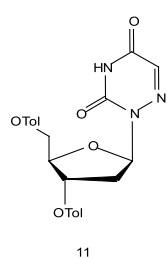
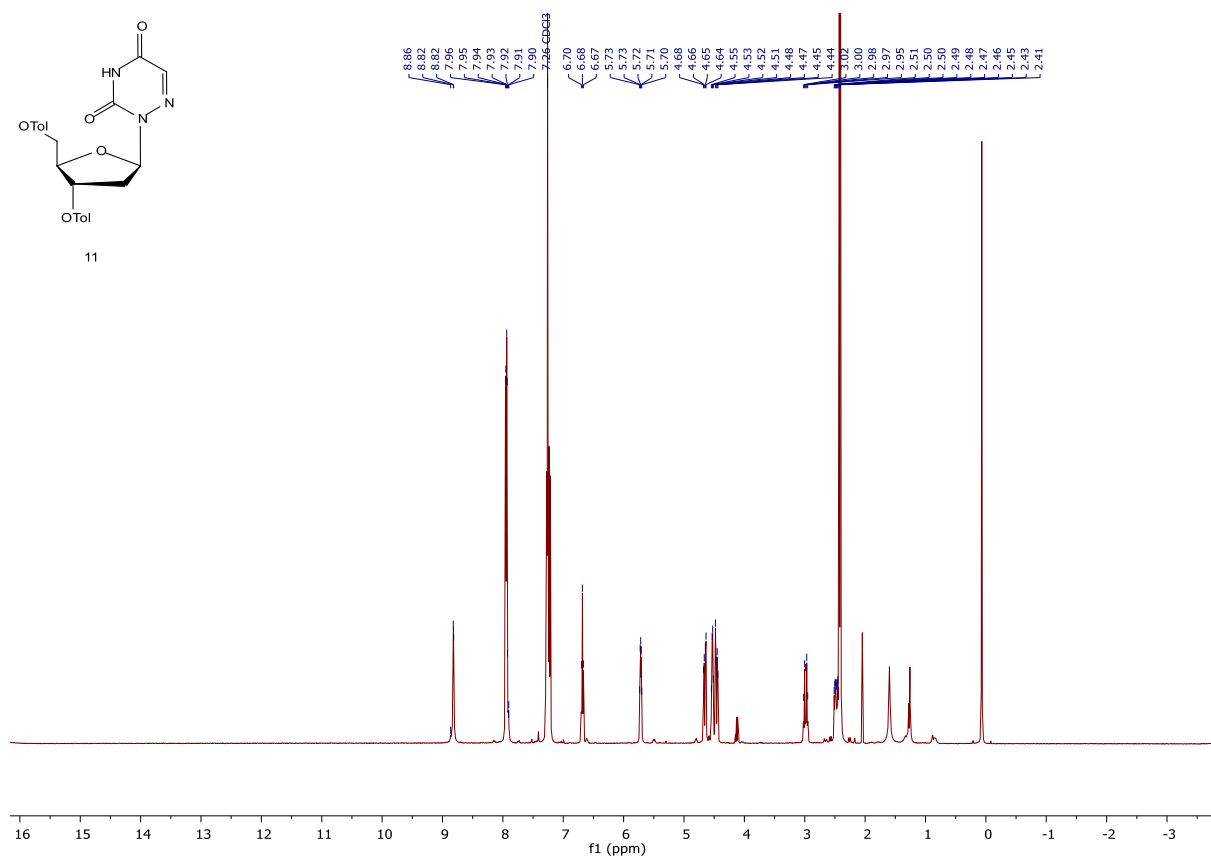
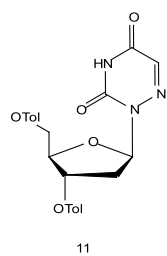


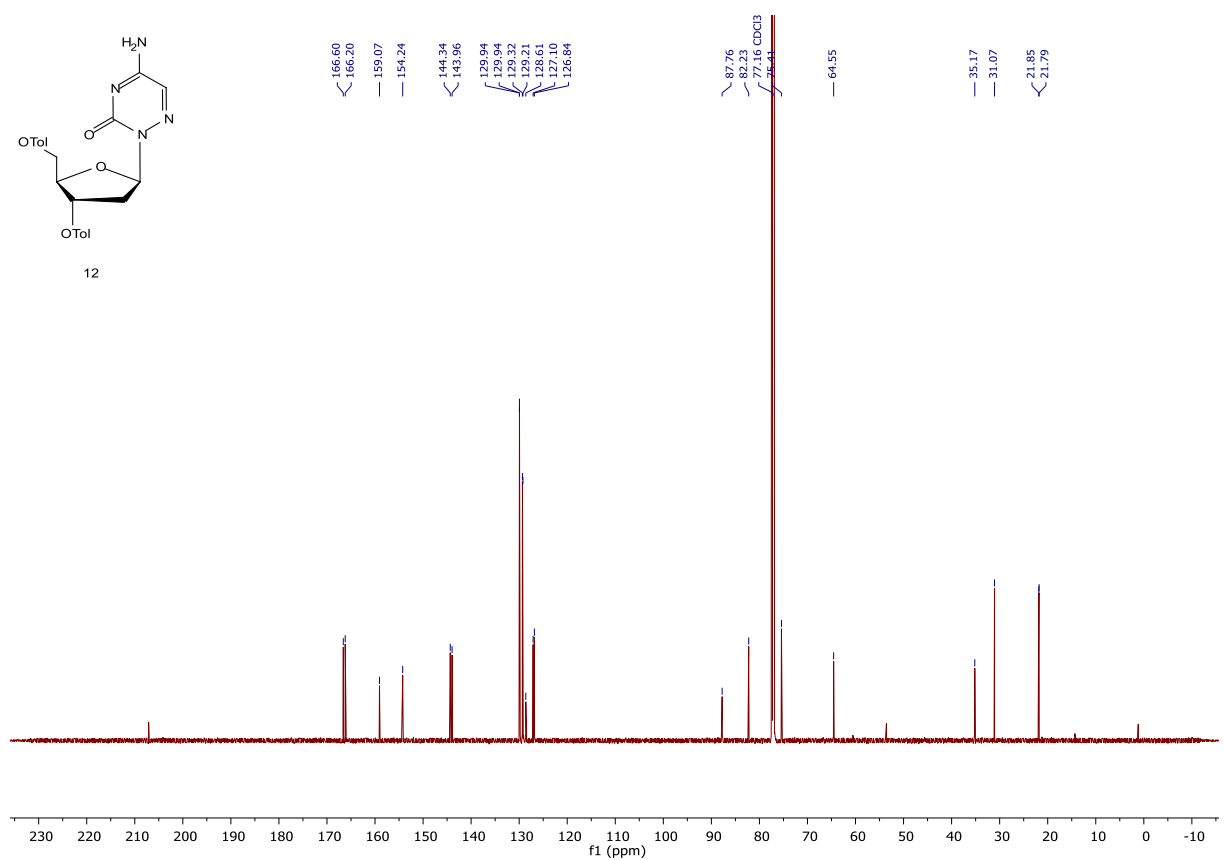


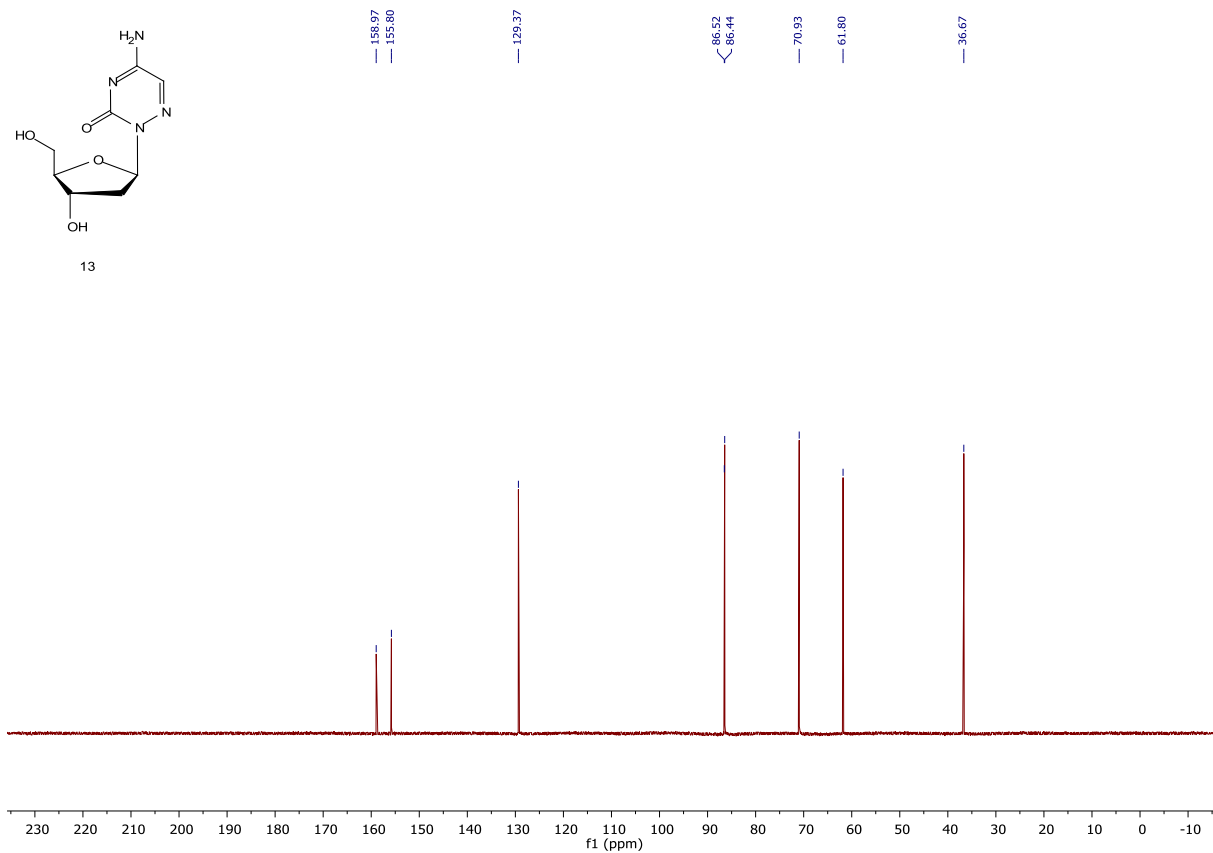
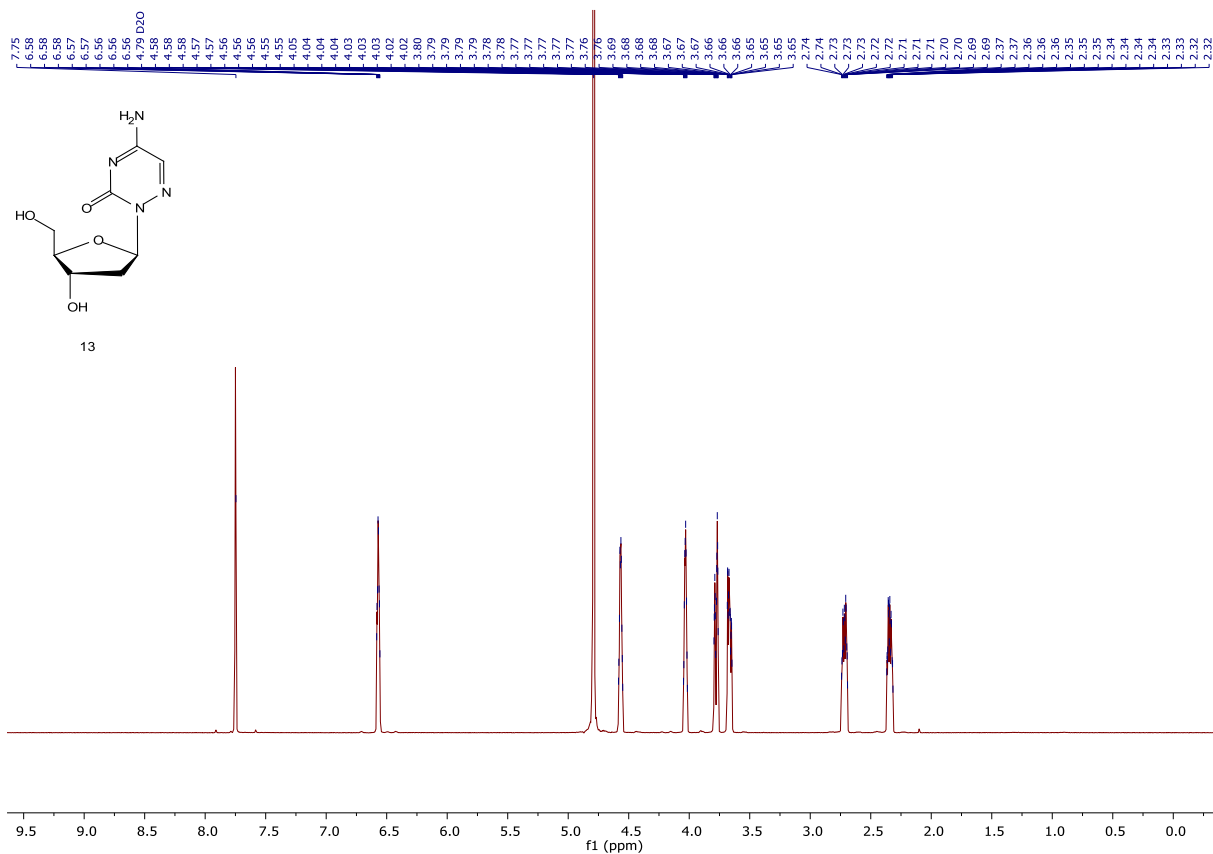


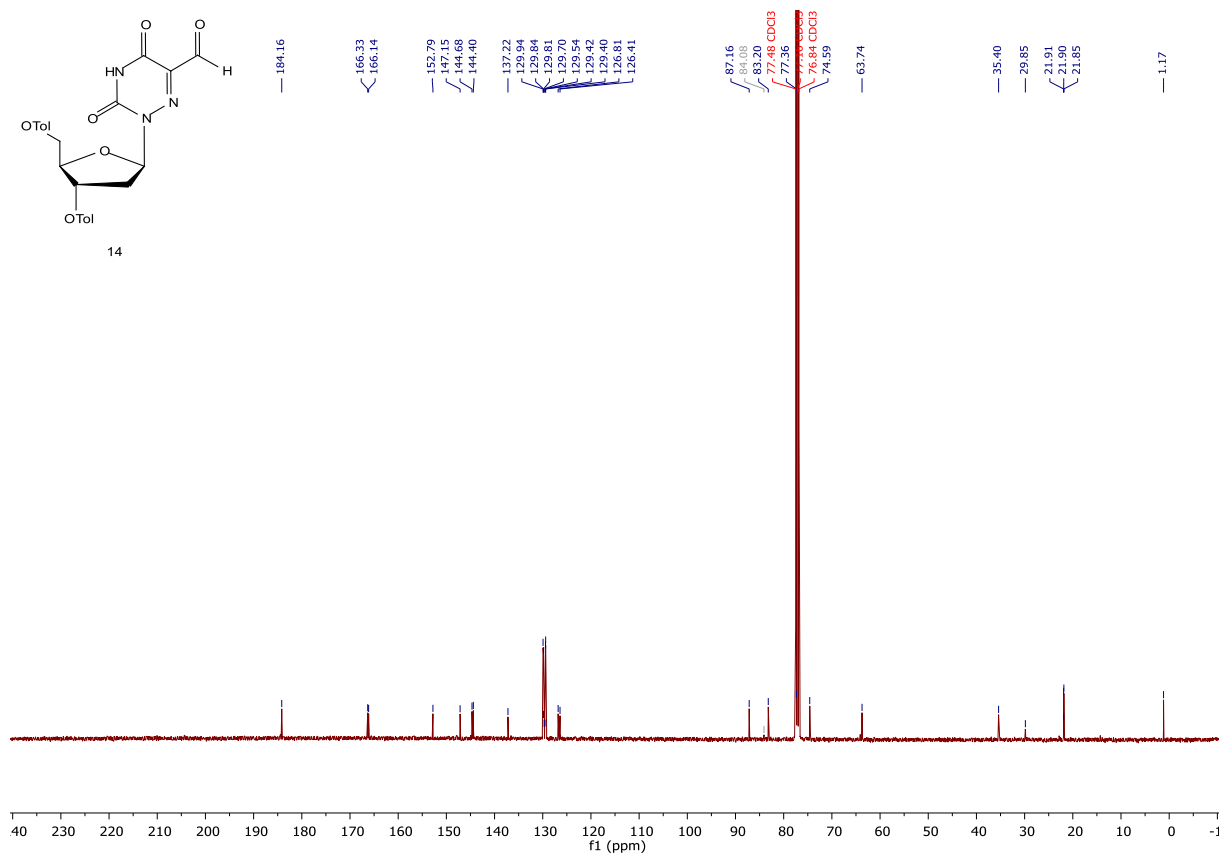
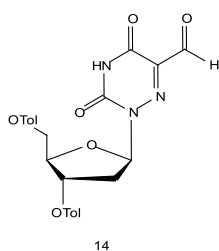
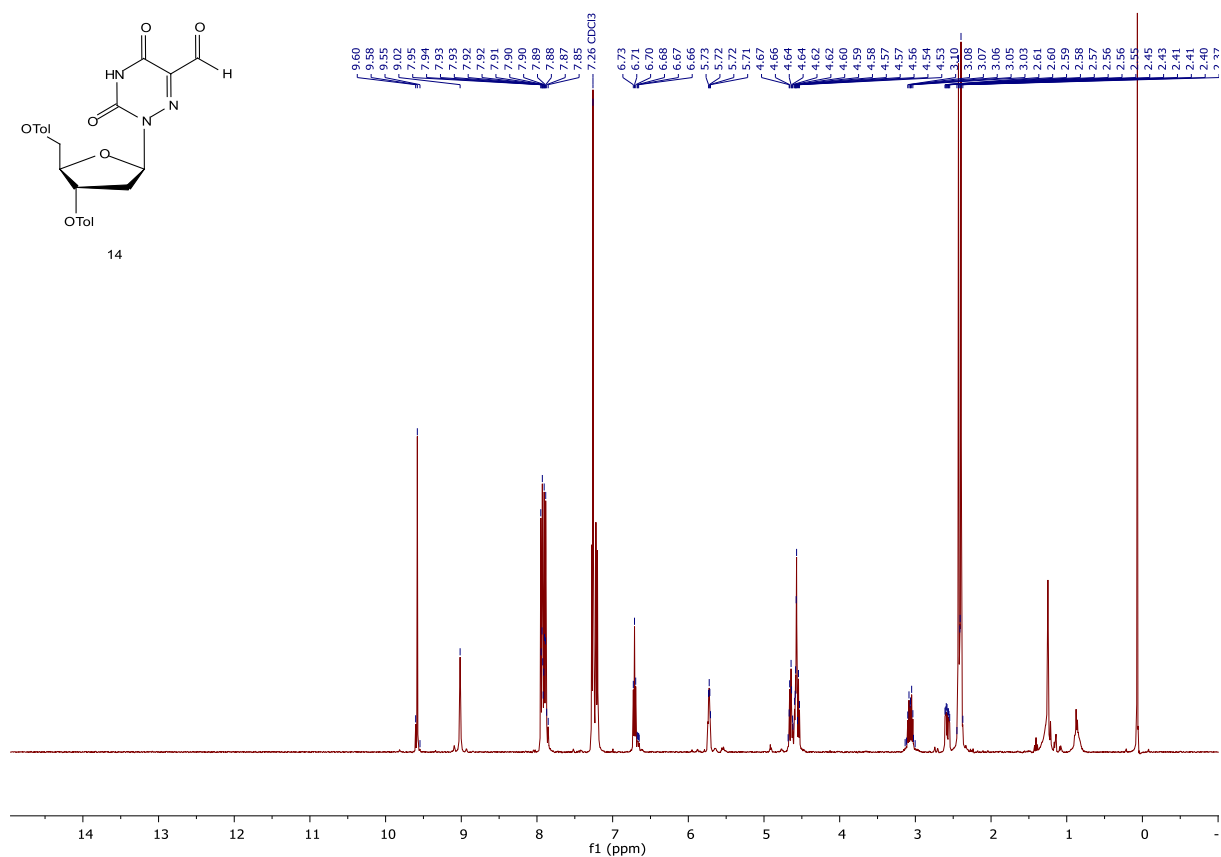
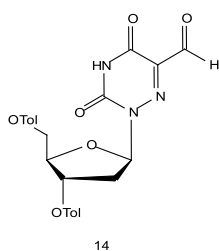


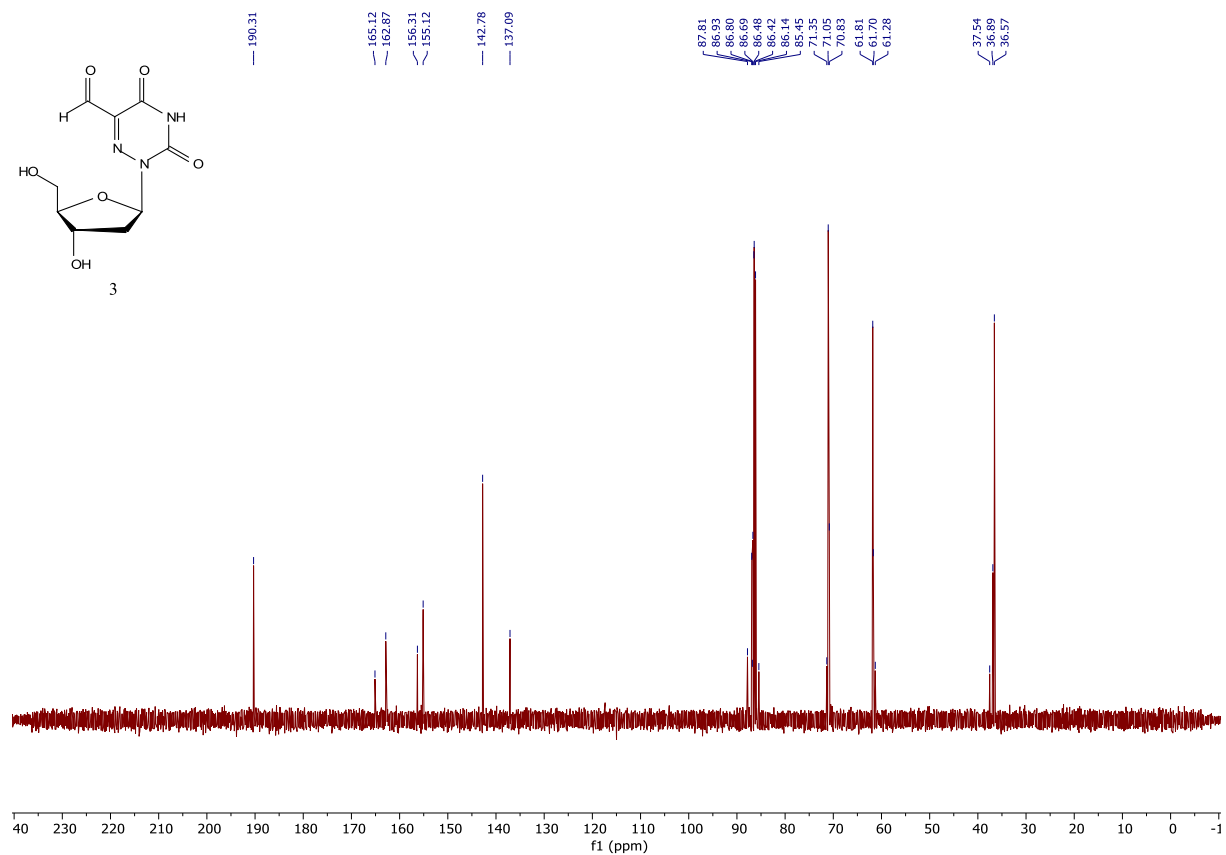
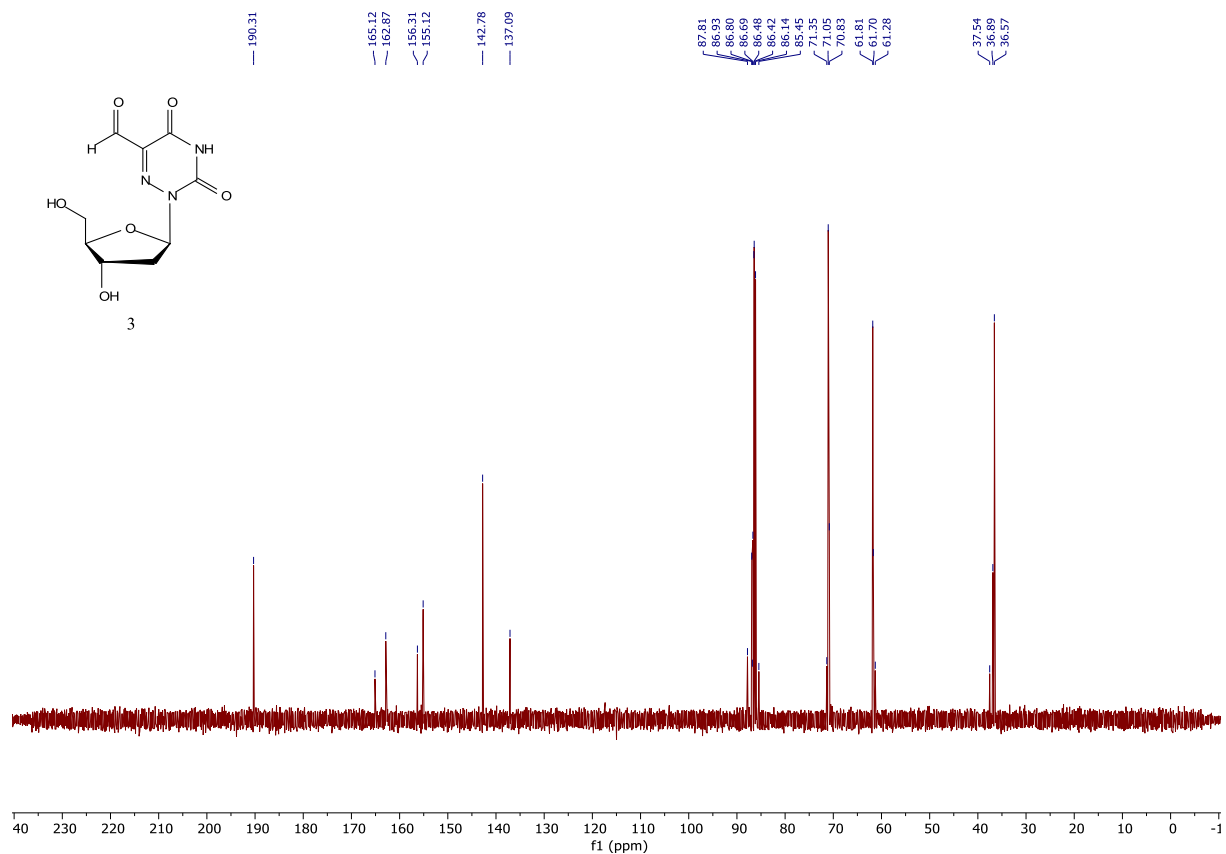












Bibliography

- [1] I. V. Alekseeva, A. S. Shalamai, V. S. Shalamai, V. P. Chernetskii, *Ukr. Khim. Zh.* **1976**, *42*, 398-401.
- [2] W. L. Mitchell, P. Ravenscroft, M. L. Hill, L. J. S. Knutsen, B. D. Judkins, R. F. Newton, D. I. C. Scopes, *J. Med. Chem.* **1986**, *29*, 809-816.
- [3] F. R. Traube, S. Schiffers, K. Iwan, S. Kellner, F. Spada, M. Müller, T. Carell, *Nat. Protoc.* **2019**, *14*, 283-312

5.2. 5-Formylcytosine enhances dsDNA kinetics by globally destabilizing the double stranded structure

Romeo C. A. Dubini, Alexander Schön, Markus Müller, Thomas Carell and Petra Rovó*

Prolog

Um den Einfluss der fdC-Modifikation auf physikalische Eigenschaften von doppelsträngiger DNA zu untersuchen, wurden im Rahmen dieser Arbeit ^1H *Chemical Exchange Saturation Transfer* (CEST) NMR-Experimente eines symmetrischen palindromischen fdC-enthaltenden DNA 12mers durchgeführt, um die Effekte der Modifikation auf atomarer Ebene zu untersuchen. Dabei konnte beobachtet werden, dass die fdC-Modifikation im untersuchten System zu veränderten Schmelzeigenschaften des Doppelstrangs bei gleichzeitig veränderten Geschwindigkeitskonstanten für Assoziation und Dissoziation des Doppelstrangsystems führte. Auf Basis der erhaltenen Ergebnisse der NMR-spektroskopischen Messungen konnten desweiteren die thermodynamischen Parameter des untersuchten Doppelstrangsystems berechnet werden.

Autorenbeitrag

Im Rahmen dieses Projekts wurden von mir die Phosphoramidite von fdC und $^{15}\text{N}_2$ -fdC synthetisiert. Das $^{15}\text{N}_2$ -fdC-enthaltende Oligonukleotid wurde allerdings nicht in der abschließenden Version der Veröffentlichung berücksichtigt. Da für die NMR-spektroskopischen Messungen große Mengen an Oligonukleotid (ca. 1 μmol DNA) benötigt wurden, führte ich die DNA-Oligonukleotidsynthese mit dem nicht-isotopenmarkierten fdC-Phosphoramidit sowie die anschließenden Entschützungen und Aufreinigungen im großen Maßstab durch. Außerdem wurden die UV/Vis-spektroskopischen Untersuchungen zum Schmelzverhalten der DNA-Oligonukleotide von mir gemessen. *Romeo Dubini* führte alle NMR-spektroskopischen Messungen mit den Oligonukleotiden durch und wertete alle erhaltenen Daten aus.

Lizenz

Kopie der *Open Access* Publikation auf Basis der *Creative Commons CC BY-NC* Lizenz, die den nicht-kommerziellen Abdruck des Artikels erlaubt.

Artikel einsehbar unter:

<https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkaa589/5870335>

Lizenz einsehbar unter:

<https://creativecommons.org/licenses/by-nc/4.0/>

Impact of 5-formylcytosine on the melting kinetics of DNA by ^1H NMR chemical exchange

Romeo C. A. Dubini^{1,2}, Alexander Schön¹, Markus Müller¹, Thomas Carell¹ and Petra Rovó^{1,2,*}

¹Faculty of Chemistry and Pharmacy, Department of Chemistry, Ludwig-Maximilians-Universität München, Butenandtstraße 5-13, 81377 Munich, Germany and ²Center for Nanoscience (CeNS), Faculty of Physics, Ludwig-Maximilians-Universität München, Schellingstraße 4, 80799 Munich, Germany

Received May 17, 2020; Revised June 24, 2020; Editorial Decision June 28, 2020; Accepted July 02, 2020

ABSTRACT

5-Formylcytosine (5fC) is a chemically edited, naturally occurring nucleobase which appears in the context of modified DNA strands. The understanding of the impact of 5fC on dsDNA physical properties is to date limited. In this work, we applied temperature-dependent ^1H Chemical Exchange Saturation Transfer (CEST) NMR experiments to non-invasively and site-specifically measure the thermodynamic and kinetic influence of formylated cytosine nucleobase on the melting process involving dsDNA. Incorporation of 5fC within symmetrically positioned CpG sites destabilizes the whole dsDNA structure—as witnessed from the $\sim 2^\circ\text{C}$ decrease in the melting temperature and $5\text{--}10\text{ kJ mol}^{-1}$ decrease in ΔG° —and affects the kinetic rates of association and dissociation. We observed an up to ~ 5 -fold enhancement of the dsDNA dissociation and an up to ~ 3 -fold reduction in ssDNA association rate constants, over multiple temperatures and for several proton reporters. Eyring and van't Hoff analysis proved that the destabilization is not localized, instead all base-pairs are affected and the transition states resembles the single-stranded conformation. These results advance our knowledge about the role of 5fC as a semi-permanent epigenetic modification and assist in the understanding of its interactions with reader proteins.

INTRODUCTION

Nucleic acids play a fundamental role within all three domains of life, however their ‘alphabet’ is intrinsically restricted to only four letters: G, C, A and T (U for RNA). In order to expand this narrow ensemble of constituents, an additional instructional layer can be integrated in the genome through chemically editing the four canoni-

cal bases, whose properties on the microscopic scale (base stacking, hydrogen bonding, solvation properties, stabilizing cross-strand terms) lie at the heart of DNA and RNA nature, structure, function and consequent biological role (1–3). Such alterations, broadly termed ‘epigenetic modifications’, have been extensively reviewed (4–7). Naturally occurring chemically edited bases are found in DNA and RNA of all kingdoms of life and play a crucial regulatory role in both development (8–12) and pathogenesis (13–17) of a given organism. Their presence influences the physico-chemical properties of the resulting nucleic acid molecule, thus reshaping the biological role of the unmodified analogs (18). Within the spectrum of epigenetic modifications in DNA, the ones involving cytosine have been studied to the largest extent, primarily due to their (i) remarkable abundance, (ii) ubiquity across all domains of life and (iii) vast range of biological roles. The most frequent chemical modification of cytosine is 5-methylcytosine (5mC), which predominantly occurs within symmetric CpG sites and is associated with stable transcriptional repression, cell differentiation, X-chromosome inactivation and imprinting (19–22). 5mC can also be further modified by Ten-Eleven Translocation (TET) enzymes, which are involved in the context of the active demethylation pathway and subsequent base excision repair (BER) cycle (23,24). The step-wise oxidation of 5mC process leads to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) in a sequential manner. In addition, protein-DNA binding structural studies have revealed that TET preferentially oxidizes 5mC in symmetric CpG sites over non-CpG sites (25). Despite being orders of magnitude less abundant than 5mC or 5hmC, 5fC is equally broadly distributed in genomic DNA (5,26). Since its discovery, 5fC has been confirmed to (i) be a genomically stable and semi-permanent modification (27,28), (ii) vary its concentration depending on the nature, health condition and developmental state of the tissue (29) and (iii) be recognized by several reader proteins with exquisite specificity over an overwhelming excess of background represented by C, 5mC, and 5hmC (30,31).

*To whom correspondence should be addressed. Tel: +49 89 2180 77652; Email: petra.rovo@lmu.de

Taken together, those evidences indicate that 5fC retains an undefined number of uncharacterized epigenetic roles, in addition to being a key intermediate in the active demethylation pathway (32).

In order to fully appreciate the nature and the extent of the various functions that 5fC performs, the properties of formylated CpG sites have been subjected to extensive investigations in the past decade. Whether 5fC is responsible for the alteration of the canonical B-DNA secondary structure is still not a firmly established notion, and contradictory conclusions have been reported (33–36). In contrast, there is reasonable consensus in accounting for the enhanced degree of dsDNA flexibility as a result of 5fC incorporation that has been attributed to an increase in the base pair wobbling around its canonical Watson–Crick, base-paired conformation (18,37,38). It has been proposed that the destabilization is due to the electron withdrawing nature of the formyl substituent, which would in turn weaken the hydrogen-bond strength between the modified base and the counter-stranded paired guanine (38,39). Recent work by Sanstead *et al.* has found that 5fC accounted for a reduction in cooperativity of duplex melting and a lowering of the dissociation barrier, providing a global characterization of heavily modified oligomers via T-jump IR spectroscopy (40). On the other hand, NMR measurement on water-exchangeable protons provided information about the extent of additional base-pair opening caused by 5fC incorporation and other modified cytosine bases in the presence of ammonia acting as a catalyst for the water-exchange process (34).

Despite the above contributions, a number of questions still remain unanswered. In particular, it is entirely unknown how 5fC affects the processes of DNA hybridization. Then, the site-specific energetics of the destabilization are unclear, i.e. whether the activation barrier between ds- to ssDNA is reduced due to a stabilization of the transition state, or due to a destabilization of the ground state. How does the transition state look like? Does it resemble the ds- or ssDNA structure? Lastly, very little is known about the range of the enhanced flexibility, i.e. how locally 5fC destabilizes dsDNA.

Characterizing the subtle differences, whether of structural or dynamical origin, of chemically modified nucleic acids molecules can be a demanding task, especially if site-specific information is needed. Solution-state NMR represents an attractive option: it does not require any bulky physical label which could confound the significance of the small chemical modification, and it allows accessing atomic resolution information on local dynamics on a wide range of time scales. NMR spin relaxation techniques have been shown to allow for such studies on a plethora of biomolecules. Out of the several time domains that can be accessed, the biologically relevant μ s-ms frame is amenable to investigation employing $R_{1\rho}$ and CPMG Relaxation Dispersion (RD) techniques (41–48). More recently another experiment, the Chemical Exchange Saturation Transfer (CEST), has been proposed (49), which allows for the elucidation of even slower events, occurring on the ms–s time scale. The essential working principle of the CEST method, whose detailed theoretical framework has been recently outlined (50,51), relies on the presence of a major, ‘ground state’ (GS) conformer of a target biomolecule that under-

goes an exchange process with one or more distinct and sparsely populated ‘excited state’ (ES) conformer(s) on the ms time scale. Experimentally, in a CEST experiment the intensity of a signal of interest is recorded as a function of the resonance offset frequency of a weakly saturating field, which is stepped through a certain interval across the region of interest. The characterization of such transiently and sparsely populated exchanging states can become invaluable in the attempt of describing the motions in both proteins and nucleic acids, yielding information on the relative populations of the involved GS and ES, the precise time frame of the exchange process, as well as an estimation of the ES chemical shift, which can be indirectly informative on the structure of the sporadic conformation (52–54).

In this study, we apply a recently described ^1H -only CEST experiment to gain insights into the exchange kinetics and thermodynamics of several proton reporters simultaneously. In this endeavour, we exploit a subset of proton probes taking advantage of the SElectively Optimized Proton Excitation (SELOPE) scheme (55), which allows for detailed NMR investigations of non-isotopically labeled samples.

To further exploit the wealth of information provided by CEST, we present an atom-specific description of the influence of 5fC base-pair thermodynamics and kinetics of both the melting and the annealing processes, which are mechanisms of primary importance given the involvement of unwound DNA in transcription, repair, and subsequent recombination. With this goal in mind, we introduce a novel methodological framework that permits the extraction of thermodynamic (ΔG° , ΔH° , ΔS° , T_m) and kinetic parameters (ΔG^\ddagger , ΔH^\ddagger , ΔS^\ddagger) from the data provided by the ^1H CEST experiments.

Our results corroborate the notion that 5fC does not perturb the canonical B-DNA conformation of dsDNA, but it does curtail the overall stability of the double-stranded structure. More specifically, it impacts the kinetic rates of both the melting and annealing processes. Combining these effects, 5fC shifts the dissociation equilibrium constant up to 9-fold towards the single-stranded conformer. By further dissection of the kinetic rates through Eyring formalism, we identified that the origin of the destabilization can be traced back to the ground state of each nucleobase involved in the dsDNA conformation, evidencing the wide range of the 5fC-induced features. In addition, we find the transition state bridging the single and double-stranded structures closely resembles the ssDNA conformation, consistent with the consensus nucleation mechanism for DNA hybridization (3).

Taken together, the above observations indicate that the repercussions arising from the inclusion of 5fC are only appreciable in a dynamic context and aid in the unravelling of its biological role. In our analysis, we provide a unified view of the processes of dissociation and association, resolving the energetics behind both processes with atomic resolution.

MATERIALS AND METHODS

Sample preparation

The canonical dsDNA sample Can₁₂ was purchased after purification via high-performance liquid chromatography

(HPLC) and desalting procedure from Ella Biotech GmbH (Martinsried, Germany).

The fdC-phosphoramidite (fdC-PA) and subsequently the modified dsDNA sample Mod₁₂ was prepared via phosphoramidite chemistry as previously reported (56). Solid phase synthesis of oligonucleotides containing fdC was performed on an ABI 394 DNA/RNA synthesizer (Applied Biosystems) using standard DNA synthesis conditions with a cartridge scale of 1 μ mol. The phosphoramidites Bz-dA, Ac-dC, iBu-dG and dT as well as the PS carriers were purchased from LinkTechnologies. For the reaction of the fdC-PA a coupling time of 180 s was applied. The terminal DMT protecting group was cleaved after DNA synthesis on the synthesizer. Basic and acidic deprotection of all oligonucleotides was performed according to literature (56). Purification of the oligonucleotides was achieved with a HPLC system (Agilent 1260 Infinity II 400 bar pump and a Agilent 1260 Infinity II VWD detecting at 260 nm) applying a buffer system of 0.1 M triethylammonium acetate in water (buffer A) and 0.1 M triethylammonium acetate in 80 % aqueous MeCN (buffer B), a gradient of 0–30% buffer B in 45 min and a flow rate of 5.0 ml/min. As stationary phase Nucleodur columns (250/10 mm, C18ec, 5 μ m) from Macherey-Nagel were used. Purified oligonucleotides were analyzed by MALDI-TOF (Bruker Autoflex II). Quantification of oligonucleotides was performed via UV/Vis spectroscopy with a NanoDrop ND-1000 spectrophotometer at 260 nm.

Both samples were dissolved in an aqueous buffer consisting of 15 mM Na₂HPO₄/NaH₂PO₄ (pH 7.0), 25 mM NaCl in H₂O. Annealing was performed by heating the dsDNA-containing buffer solution to 90°C for 5 min and slowly cooling it to 5°C in ~90 min, after which it was allowed to return to room temperature. Then, the NMR sample was prepared with the addition of 0.02% NaN₃, 25 μ M DSS and 5% D₂O, resulting in final sample concentrations of 1.00 and 1.58 mM for Can₁₂ and Mod₁₂, respectively, as determined via UV spectrophotometric measurements at 260 nm using the extinction coefficient calculated via the nearest neighbor approximation.

UV/Vis spectroscopy

UV/Vis melting profiles of the oligonucleotides were measured at 260 nm with a JASCO V-650 UV/Vis spectrophotometer between 20 and 85°C (scanning rate of 1°C/min), and each sample was measured four times. Samples were placed into 100 μ l cuvettes and diluted with the same Na₂HPO₄/NaH₂PO₄, NaCl aqueous buffer as used in the NMR experiment. Before each measurement, a layer of mineral oil was placed on the surface of the sample in order to prevent water evaporation. Mod₁₂ was measured at six concentrations (1.23, 1.53, 2.45, 3.72, 6.14 and 12.27 μ M), Can₁₂ at four (1.17, 1.76, 3.52 and 7.04 μ M). All concentration values yielded absorption values within the linear range of the spectrometer.

NMR spectroscopy

All experiments were performed on Bruker Avance III spectrometer operating at a ¹H Larmor frequency of 800 MHz

(corresponding to a magnetic field of 18.8 T) equipped with a 5 mm triple-resonance cryogenically cooled probe. Standard 2D NOESY (mixing time of 150–250 ms, shown in Supplementary Figure S1D and E) spectra were recorded at 37°C for resonance assignment. Site-selective spin relaxation measurements were performed following the SELOPE scheme; these included ¹H CEST, on-resonance ¹H R_{1 ρ} , recorded with a spin-lock strength of 4 kHz, and ¹H R₁ experiments at temperatures between 37 and 62°C. The employed R_{1 ρ} and ¹H CEST pulse sequences have been modified from Schlagnitweit *et al.* (55) in order to allow for data acquisition in an pseudo-3D interleaved fashion, while an in-house pseudo-2D ¹H R₁ pulse sequence (displayed in Supplementary Figure S5) with selective excitation and water suppression was used to record ¹H inversion recovery. ¹H R_{1 ρ} relaxation decays were collected with time delays of 20, 30, 40, 60, 80, 100 and 120 ms, while the inversion recovery profile of the ¹H R₁ measurement was collected with time delays of 0.5, 1, 5, 10, 25, 50, 75, 100, 200, 500, 750 and 1000 ms. All ¹H CEST profiles were recorded with a spin lock B₁ field of 25 Hz using a T_{EX} of 0.4 s, additional measurement conditions for each temperature are summarized in Supplementary Table S1. All NMR experiments were recorded with a relaxation delay (*d*₁) of 1 s, with the exception of R₁ measurements for which *d*₁ was set to 10 s. When not at the water resonance to allow for water suppression, the ¹H carrier was positioned in the middle of the H8/H6/H2 region allowing for selective irradiation of the desired aromatic sites (Figure 1). The CEST measurements were performed at 37, 45, 53, 57, 58, 59, 60, 61, 62°C for Can₁₂ and at 37, 45, 53, 55, 56, 57, 58, 59, 60°C for Mod₁₂. In total, eleven sites for Can₁₂ and twelve for Mod₁₂ (namely for C2H6, G3H8, A4H2, A4H8, T5H6, C6H6, A8H2, A8H8, T9H6, C10H6, and C12H6 for Can₁₂ and G1H8, C2H6, A4H2, A4H8, T5H6, C6H6, C6H7, A8H2, A8H8, T9H6, C10H6, and C12H6 for Mod₁₂) could be uniquely picked and assigned in the CEST spectra. Reconstruction of pseudo-3D interleaved spectra was achieved through an in-house Perl script. Both assignment and spin relaxation data was processed and inspected using TopSpin 3 or 4. Calibration of the DSS signal at 0 ppm allowed for exact spectral referencing at all temperatures. Subsequent ¹H resonance assignments and raw relaxation data extraction (including peak height and signal to noise) were carried out in Sparky (57).

RESULTS

We studied the following 12mer, homo-formylated palindromic sequence: 5'-GCGAT-X-GATCGC-3', where X = 5fC in the modified sample, Mod₁₂ and unedited C in the canonical sample, Can₁₂ (Figure 1A). The selection was made taking into consideration that (i) it contains a centrally positioned CpG domain, which is the naturally occurring settings of 5fC and (ii) being a self-complementary sequence, the resulting ¹H-based NMR spectra are less crowded, and thus less prone to be affected by overlapping signals. The latter feature can be ascertained by inspecting the 2D SELOPE spectrum of Mod₁₂ (Figure 1B), where each signal arising from the H6 of C and T, and H8, H2 protons of A in Mod₁₂ is resolved. Guanine H8 protons

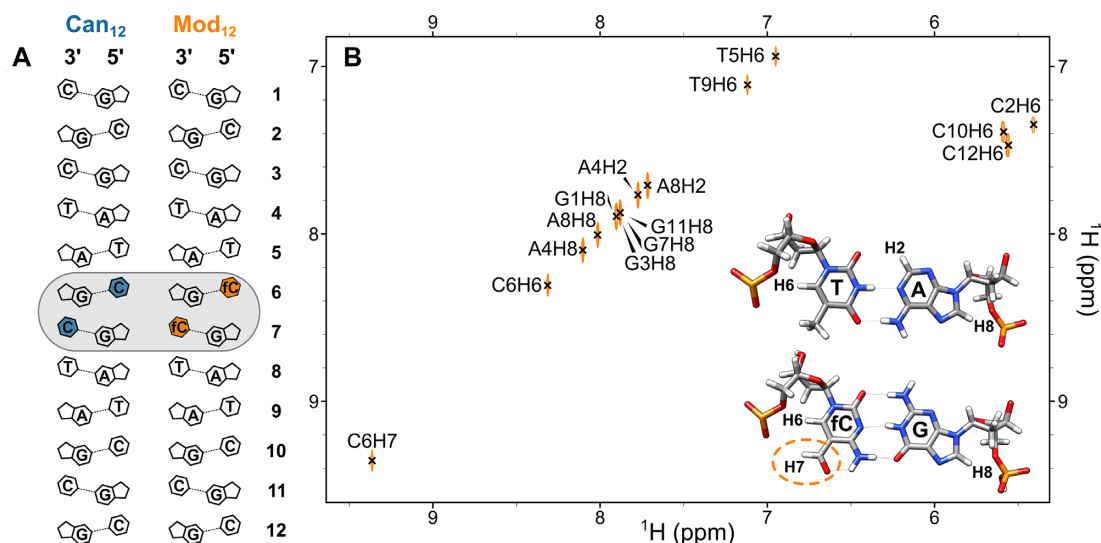


Figure 1. (A) DNA constructs of Can₁₂ and Mod₁₂, gray box highlights the site of the 5fC incorporation. (B) SELOPE 2D ¹H-¹H correlation spectrum of the Mod₁₂ aromatic region. H6 protons of canonical cytosine bases appear as crosspeaks due to their ³*J* coupling to H5, while the H6 proton of the formylated cytosine C6, lacking the coupling partner H5, appears at the diagonal. The formyl proton C6H7 appears at 9.4 ppm.

represent an exception as they almost completely overlap for G3, G7 and G11 H8 in Mod₁₂ in CEST spectra, thus no site specific analysis was possible for these sites.

Structural impact

A set of NOESY and natural abundance ¹H-¹³C HSQC spectra were recorded to weigh on the influence of the 5fC modification on the ground-state conformation of DNA. The very weak intensity of the H2'-H4' cross-peaks in NOESY spectra (indicative of the C2'-endo sugar pucker) and dihedral angles obtained from the ³*J*¹H-¹H couplings of the sugar moieties confirmed the predominant presence of B-DNA conformation in both samples (58).

The chemical shift values for all ¹H and ¹³C nuclei are highly comparable between the two samples, as expected, the only differences occur for sites nearby the modified base. Most notably, the resonance frequency of both amino protons of the formylated base shift markedly downfield from 6.5 and 8.2 to 7.9 and 9.1 ppm, respectively indicating the formation of a strong hydrogen bond between the formyl group and the amino protons, as per the predicted 5fC:G base-pair in Figure 1 B). This intramolecular hydrogen bond is stable even at high temperatures, only marginally weakening (0.08 ppm upfield shift) when the sample approaches the melting temperature. Besides, the 0.4 ppm upfield shift the G7 imino proton resonance of the modified sample with respect to the same proton of Can₁₂ (Supplementary Figure S1A-C) suggests a reduced H-bonding and thus weakened base-pairing capacity in the 5fC:G with respect to C:G pair as observed by others (34,38).

Aldehydes in aqueous solution can undergo an acid/base catalyzed hydration reactions (59). Contrary to previous reports on different dsDNA sequences featuring 5fC in the context of X:C mismatches (where X is either 5mC, 5hmC, 5fC or 5caC) (60), we do not detect resonances compatible with such a hydrate form neither in ¹H nor in ¹³C based experiments.

In accordance with previous studies, we did not identify any difference in the ground-state conformation of Can₁₂ and Mod₁₂ (34-36). The two samples displayed localized and discrete chemical shifts in both the ¹H and ¹³C dimensions due to the presence of the formyl substituent on the edited cytosine (C6) ring and the nearest neighbouring bases T5 and G7 (Supplementary Figure S1). Considering the affinities between the spectral features of the two samples, we conclude that 5fC is well tolerated in the context of B-DNA conformation.

Concentration dependent UV/Vis melting studies

As an initial assessment of the effect of 5fC on the stability of the dsDNA conformer, we determined the concentration-dependent melting temperature (*T*_m) of the two samples via UV/Vis spectroscopy. Since DNA annealing is a bimolecular process, the melting temperature of the double-stranded conformation is highly concentration dependent, causing its *T*_m to steadily increase with the molarity of the sample (61). Such relationship can be exploited to obtain an overall estimation about the melting thermodynamics (ΔG° , ΔH° , ΔS°) and to extrapolate to the expected melting temperature to higher concentrations; in our case to the concentration range of the NMR samples which is approximately two- to three-orders-of-magnitude higher than that of the UV/Vis samples.

Hence, we recorded the melting profiles of both samples at multiple concentrations in the 1-10 μ M range, and extracted the thermodynamic parameters assuming a linear relationship between $\ln(C_t)$ and *T*_m, where *C*_t is the total monomer concentration (Supplementary Figure 2, S2 and S3). This method assumes that ΔH° is temperature independent i.e. the specific heat capacity (*C*_p) is constant within the temperature range under consideration. In Table 1, we compare the fitted bulk thermodynamic parameters with the site-specific thermodynamic parameters obtained from CEST NMR measurements (*vide infra*). Systematic devia-

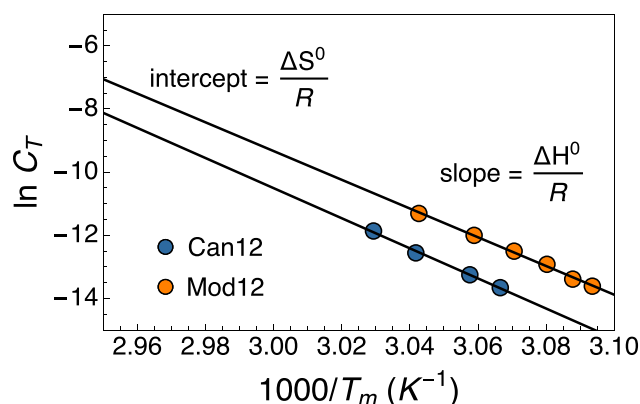


Figure 2. Concentration-dependent UV/Vis melting temperatures analysis of Can₁₂ (blue) and Mod₁₂ (orange). The reciprocal of the observed T_m is plotted as a function of the logarithm of the total DNA concentration (C_T). The slope and the intercept of the fitted linear corresponds to $\Delta H^\circ/R$, and $\Delta S^\circ/R$, respectively. The obtained thermodynamic parameters and the expected T_m at 1 mM concentration are listed in Table 1.

tion between the derived thermodynamic parameters of the two methods indicate a non-zero ΔC_p , which is expected but rarely taken into account explicitly (62).

The fitted standard enthalpy and entropy changes, and derived Gibbs free energy differences indicate that inclusion of a 5fC site in a dsDNA oligomer leads to a consistent decrease in stability. The lower extent of cohesion between the two strands in Mod₁₂ is apparent from the decrease of both ΔH° and ΔS° with respect to ΔH° and ΔS° of Can₁₂. The enthalpy–entropy compensation is only partial, leading to a decrease of ΔG° by 5 kJ mol^{−1} at 37°C and to a lowered melting temperature ($\Delta T_m = 2^\circ\text{C}$) at 1 mM total monomer concentration.

Temperature-dependent CEST profiles

While UV/Vis melting studies provided a coarse glimpse into the impact of 5fC modification on the overall DNA stability, they are not sufficient to account for the kinetics related to the hybridization and melting processes. Conversely, CEST measurements recorded in a temperature-dependent manner can reveal both the kinetics and thermodynamics of ssDNA association and dsDNA dissociation in a site-specific manner. Therefore, we directed our efforts to study the melting profiles of Can₁₂ and Mod₁₂ using ¹H CEST experiments.

Overall, we recorded CEST profiles in a SELOPE pseudo-3D interleaved fashion for both samples at multiple temperatures between 37 and 62°C. We focused on the spectral region between 6.5 and 10 ppm, populated by resonances belonging to purines' H8 and H2 and pyrimidines' H6 in addition to the formyl proton of the formylated cytosine site (C6H7) (Figure 1B).

Due to reliance of the CEST technique on the non-equal chemical shifts between the interconverting states (in this case between dsDNA and ssDNA), only those sites could be analysed where this condition was met. The population of the ssDNA species fell below the detection limit at the three lowest measured temperatures (37, 45 and 53°C for both

samples), and hence, those data sets were not employed in the parameter fitting procedure (see Supplementary Data text for details about the fitting and error estimation). High-temperature CEST profiles together with their fits to a two-state exchange model for three exemplary proton reporters (T5H6, C6H6 and A8H8) are displayed in Figure 3, and the rest is shown in Supplementary Figure S4.

We observed a single excited state species across all conditions for both samples. Coincidentally, the chemical shift of the single, uniquely detectable ES closely resembles the predicted value for a random coil shifts, whose values were calculated considering the nearest and next-nearest neighbor effect (63,64). A comparison between the CEST-derived chemical shifts and the predicted values is offered in Supplementary Table S2. This outcome suggests that there is a single predominant exchange process which involves melting of the dsDNA and annealing of ssDNA, and no detectable bulge or kink formation is present on the slow ms timescale. A comparison between the proton chemical shift values of the ds- and ssDNA strands is shown in Figure 4, which displays the simulated 1D spectra featuring the sites that could be identified and characterized via CEST. The simulations used the chemical shifts and R_1 , R_2 relaxation rate constants that were obtained from the CEST fitting procedure.

We proceeded by fitting the CEST profiles assuming an all-or-none two-state exchange model (dsDNA \rightleftharpoons 2 ssDNA), which we consider to be a reasonable assumption given the relatively short length of the selected sequence (65). As fitted parameters we obtained populations (p_b for dsDNA, $1 - p_b$ for ssDNA), exchange kinetics (k_{ex}), and chemical shifts of the exchanging states at each of the five highest temperatures, while we used the the sixth (lowest temperature of the ensemble for each sample) CEST profile to compare the predicted back-calculated value from the fits to the experimental value. To ease the fitting procedure, we measured ¹H longitudinal (R_1) and transverse (R_2) relaxation rates separately at multiple temperatures and used those rates as inputs for the CEST fits assuming that the rates of the dsDNA and ssDNA states are the same. The details of the fitting procedure is discussed in the Supplementary Data Text, and the fitted values are listed in Supplementary Tables S3 and S4.

Site specific quantification of thermodynamic and kinetic parameters

To obtain a rigorous, site-specific assessment of the DNA melting and annealing processes and the associated energetics, we analysed the CEST-derived site specific populations and exchange rates in the context of Eyring and van't Hoff theories. Within the proposed framework, concentrations are explicitly taken into account and extracted parameters are thus normalized and amenable to a scrupulous comparison.

The observed exchange rate in the CEST experiment (k_{ex}) is the sum of the forward (k_1) and backward (k_{-1}) reactions, where the forward rate is equivalent to the rate of duplex dissociation ($k_1 = k_d$), and the backward rate is related to the association rate, k_a

$$k_{-1} = 2[\text{ssDNA}]_{eq}k_a, \quad (1)$$

Table 1. Thermodynamic and kinetic parameters of the dsDNA melting process obtained from the van't Hoff and Eyring analysis of the CEST-derived exchange parameters, and from UV/Vis melting studies. Errors are given as one standard deviation

		Thermodynamics		Dissociation kinetics					Association kinetics		
		ΔG_{37C}^0 kJ mol ⁻¹	ΔH^0 kJ mol ⁻¹	ΔS^0 J K ⁻¹ mol ⁻¹	T_m (1 mM) °C	$\Delta G_{d,37C}^\ddagger$ kJ mol ⁻¹	ΔH_d^\ddagger kJ mol ⁻¹	ΔS_d^\ddagger J K ⁻¹ mol ⁻¹	$\Delta G_{a,37C}^\ddagger$ kJ mol ⁻¹	ΔH_a^\ddagger kJ mol ⁻¹	ΔS_a^\ddagger J K ⁻¹ mol ⁻¹
Mod ₁₂	G1H8	64.6 ± 7.4	470 ± 128	1306 ± 389	75 ± 14.1	86.8 ± 4.6	122 ± 79	115 ± 241	22.2 ± 4.7	-347 ± 80	-1191 ± 244
	C2H6	74.7 ± 10.4	769 ± 158	2238 ± 475	62 ± 0.5	99.7 ± 5.6	413 ± 84	1011 ± 253	25. ± 5.5	-356 ± 83	-1227 ± 250
	A4H2	73.1 ± 1.	712 ± 16	2060 ± 49	63 ± 0.1	101.6 ± 0.6	432 ± 9	1064 ± 28	28.5 ± 0.6	-281 ± 9	-996 ± 28
	A4H8	71.9 ± 1.2	716 ± 19	2078 ± 58	62 ± 0.1	101.7 ± 1.2	444 ± 18	1102 ± 55	29.8 ± 1.1	-273 ± 18	-975 ± 53
	T5H6	64.3 ± 1.1	613 ± 18	1768 ± 53	62 ± 0.1	93.9 ± 0.8	340 ± 12	792 ± 38	29.6 ± 0.8	-273 ± 12	-976 ± 37
	C6H6	67.4 ± 1.3	652 ± 20	1886 ± 60	63 ± 0.1	99.9 ± 1.6	422 ± 24	1038 ± 73	32.5 ± 1.2	-230 ± 19	-848 ± 58
	C6H7	48.6 ± 1.6	344 ± 26	953 ± 78	68 ± 0.8	118.8 ± 1.8	640 ± 28	1680 ± 86	70.2 ± 1.9	296 ± 30	727 ± 90
	A8H2	71.2 ± 0.6	678 ± 9	1955 ± 28	64 ± 0.1	99.6 ± 0.4	398 ± 6	962 ± 18	28.4 ± 0.4	-280 ± 6	-994 ± 18
	A8H8	65.3 ± 0.9	604 ± 14	1737 ± 42	63 ± 0.1	98.2 ± 0.7	383 ± 11	920 ± 34	32.9 ± 0.6	-221 ± 10	-817 ± 31
	T9H6	64.6 ± 1.2	611 ± 18	1761 ± 55	63 ± 0.1	97.1 ± 1.	383 ± 16	920 ± 47	32.5 ± 1.	-228 ± 15	-841 ± 45
	C10H6	66.4 ± 1.2	636 ± 19	1835 ± 57	63 ± 0.2	97.5 ± 1.	384 ± 16	924 ± 49	31.1 ± 1.1	-252 ± 17	-911 ± 50
	C12H6	69.7 ± 1.	679 ± 15	1964 ± 47	63 ± 0.1	98.7 ± 0.9	402 ± 14	977 ± 43	29. ± 1.	-277 ± 15	-987 ± 45
	UV/Vis	50.4 ± 1.2	338 ± 25	1056 ± 77	66 ± 1.0						
	C2H6	71.1 ± 1.7	641 ± 26	1838 ± 77	65 ± 0.2	99.2 ± 2.2	372 ± 31	880 ± 94	28.1 ± 2.1	-269 ± 30	-958 ± 91
Can ₁₂	G3H8	69. ± 1.3	608 ± 20	1738 ± 59	66 ± 0.2	106.4 ± 1.9	459 ± 28	1138 ± 84	37.4 ± 1.6	-149 ± 23	-600 ± 71
	A4H2	72.5 ± 0.9	626 ± 14	1784 ± 41	67 ± 0.2	100.1 ± 0.7	363 ± 11	847 ± 32	27.7 ± 0.7	-263 ± 11	-936 ± 32
	A4H8	77.6 ± 1.1	732 ± 16	2110 ± 48	65 ± 0.1	108.3 ± 1.8	498 ± 25	1258 ± 76	30.7 ± 1.8	-234 ± 25	-852 ± 75
	T5H6	71. ± 1.6	633 ± 23	1812 ± 70	65 ± 0.2	102.7 ± 1.4	415 ± 20	1006 ± 60	31.7 ± 1.3	-218 ± 19	-806 ± 57
	C6H6	70.7 ± 1.4	638 ± 21	1829 ± 63	65 ± 0.2	99.8 ± 1.2	379 ± 17	900 ± 51	29.1 ± 1.2	-259 ± 17	-929 ± 51
	A8H2	73.8 ± 1.1	643 ± 17	1834 ± 50	67 ± 0.2	103.6 ± 1.	409 ± 15	986 ± 45	29.8 ± 1.	-233 ± 14	-849 ± 42
	A8H8	73.1 ± 1.	666 ± 15	1910 ± 45	65 ± 0.1	104.3 ± 1.	439 ± 14	1078 ± 43	31.2 ± 0.9	-227 ± 13	-833 ± 40
	T9H6	70.2 ± 1.8	626 ± 27	1791 ± 80	65 ± 0.3	102.2 ± 1.5	410 ± 22	993 ± 66	31.9 ± 1.5	-216 ± 22	-798 ± 65
	C10H6	71.3 ± 1.4	642 ± 21	1841 ± 64	65 ± 0.2	100.3 ± 1.3	382 ± 19	908 ± 56	29. ± 1.3	-260 ± 20	-933 ± 59
	C12H6	77.6 ± 1.1	723 ± 17	2082 ± 50	65 ± 0.1	104.2 ± 1.3	434 ± 18	1064 ± 55	26.6 ± 1.3	-289 ± 19	-1018 ± 58
	UV/Vis	55.2 ± 2.5	408 ± 43	1130 ± 131	69 ± 1.5						

where $[\text{ssDNA}]_{\text{eq}} = C_t(1 - p_b)$ is the equilibrium concentration of the ssDNA, and C_t is the total ssDNA concentration. The expression of equation (1) was derived assuming a weak perturbation as described in the Supplementary Data (66). k_a and k_d kinetic rates can be inferred from fitted k_{ex} and p_b parameters as follows:

$$k_a = \frac{k_{\text{ex}} p_b}{2C_t(1 - p_b)} \quad (2)$$

$$k_d = k_{\text{ex}}(1 - p_b) \quad (3)$$

The temperature dependence of k_d and k_a can be used to derive the activation barriers for the dissociation ($\Delta G_d^\ddagger(T)$) and association ($\Delta G_a^\ddagger(T)$) processes. Subsequently, the corresponding activation enthalpy (ΔH_d^\ddagger , ΔH_a^\ddagger) and entropy (ΔS_d^\ddagger , ΔS_a^\ddagger) changes can be obtained. Further, the ratio between $k_d(T)$ and $k_a(T)$ rates, which gives the equilibrium constant ($K_d(T)$) for dissociation, is relevant for the calculation of the site-specific melting thermodynamics parameters (ΔG° , ΔH° , ΔS° and T_m).

In order to infer on the thermodynamics of the process, we assumed ΔS° and ΔH° to be temperature independent (so-called van't Hoff assumption) and fitted the variation of K_d with temperature accordingly

$$K_d(T) = \exp\left(\frac{-\Delta H^\circ}{RT} + \frac{\Delta S^\circ}{R}\right) \quad (4)$$

leading to a linear dependence between $\ln K_d$ and $1/T$ (Figure 6). See Supplementary Data text for the details of the derivation of the thermodynamic parameters.

Consistently with published experimental and computational DNA melting kinetic studies, we find that the

temperature-dependence of the k_d (k_a) rates follow a Eyring-Polanyi relationship and they increase (decrease) with increasing temperature (3,65). The dissociation is a five orders-of-magnitude slower process than the bimolecular association, the former being in the range of 0.5–10 s⁻¹, and the latter in 0.5–2 × 10⁵ M⁻¹ s⁻¹. When comparing the two samples at the same temperatures (Figure 5), we find that the k_d rates are consistently and systematically 2- to 5-fold higher for Mod₁₂ than for Can₁₂.

In analogy to the dissociation rates, also k_a values show a regular pattern. We detect the hybridization process to be 1.5- to 3.5-fold faster in the canonical sample. These results indicate that 5fC simultaneously facilitates the melting process and hinders the annealing process.

Figure 6 shows the logarithm of the obtained kinetic rates and equilibrium constants for T5H6, C6H6 and A8H8 protons as a function of $1000(1/T - 1/T_{\text{hm}})$, where $T_{\text{hm}} = n(\sum_i T_i^{-1})^{-1}$ is the harmonic mean of the temperatures used in the CEST studies, which is 58 and 60°C for Mod₁₂ and Can₁₂, respectively. The linear change of the kinetic rates with T^{-1} suggests a single, well-defined transition state. According to a modified version of the Eyring-Polanyi equation, which takes into account the statistical bias induced by the different temperatures used, the dissociation and association rate can be given as

$$k_i(T) = \frac{\kappa k_B T}{h} \times \exp\left(-\frac{\Delta G_i^\ddagger(T_{\text{hm}})}{RT_{\text{hm}}} - \frac{\Delta H_i^\ddagger}{R} \left(\frac{1}{T} - \frac{1}{T_{\text{hm}}}\right)\right) \quad (5)$$

where i stands for association (a) or dissociation (d), k_B is the Boltzmann constant, R is the universal gas constant,

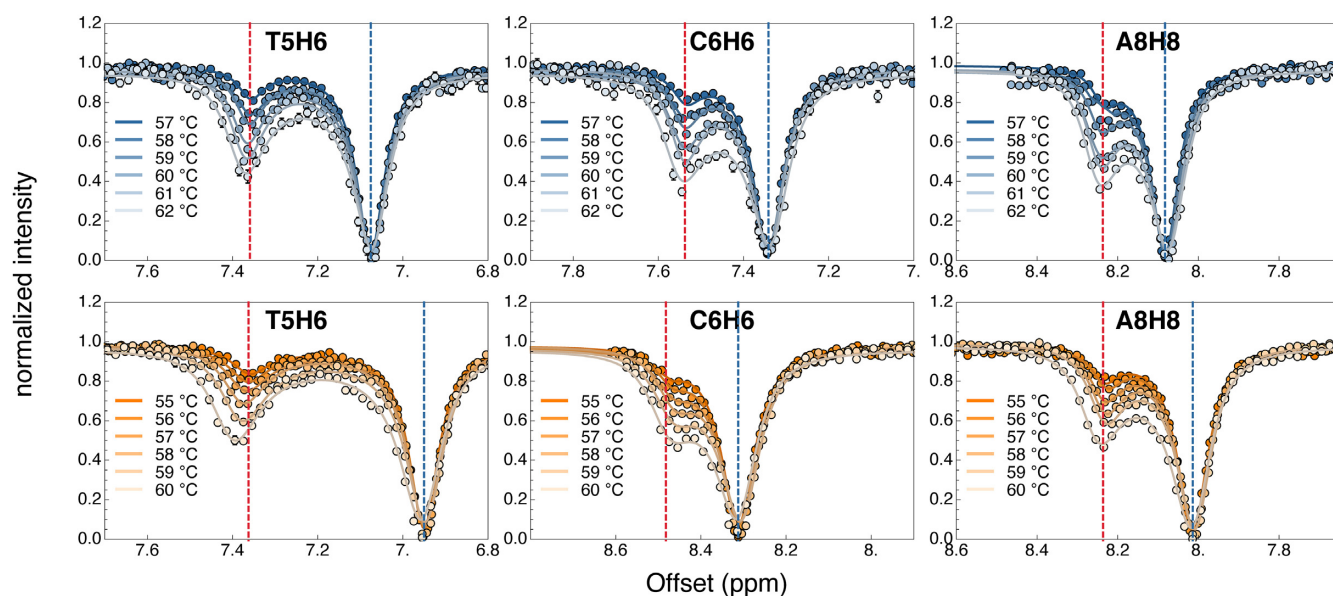


Figure 3. Temperature-dependent ^1H CEST profiles and fits of T5H6, C6H6 and A8H8 protons of Can₁₂ (blue, upper row) and Mod₁₂ (orange, bottom row). Dashed blue and red vertical lines indicate the chemical shifts of the double-stranded and single-stranded conformations, respectively. As expected, the population of ssDNA increases gradually with increasing temperature.

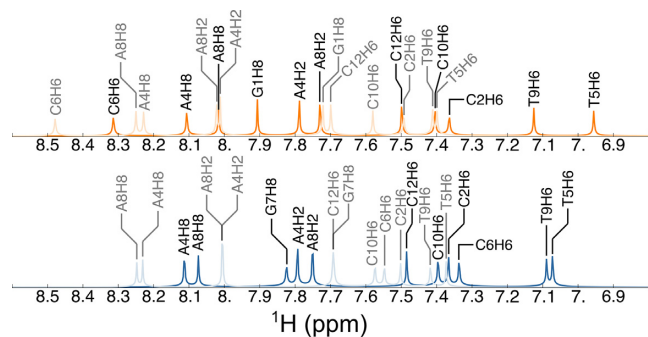


Figure 4. Simulated ^1H 1D spectra. Darker and lighter shades represent the ground-state (dsDNA) and excited-state (ssDNA) conformations. Only those sites are included that showed a second dip in the CEST profiles.

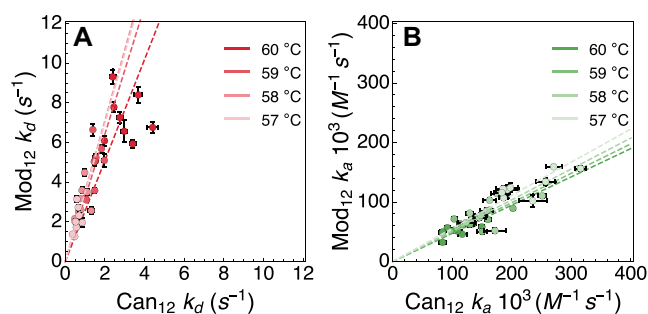


Figure 5. Correlation between the site-specific k_d (A) and k_a (B) kinetic rates between the observable sites of Mod₁₂ and Can₁₂ at 57, 58, 59 and 60°C. Dashed lines represent the linear fits at each temperature with a slope of 2.5–3.5 for k_d and 0.47–0.55 for k_a .

and κ is the transmission coefficient, assumed to be 1. The slope and intercept of the linear fit to $\ln k_i/T$ versus $1000(1/T - 1/T_{\text{hm}})$ gives the activation enthalpy change

(ΔH_i^\ddagger) and the Gibbs free energy change at T_{hm} ($\Delta G_i^\ddagger(T_{\text{hm}})$), respectively, from which the activation entropy (ΔS_i^\ddagger) can be derived as

$$\Delta S_i^\ddagger = \frac{\Delta H_i^\ddagger - \Delta G_i^\ddagger(T_{\text{hm}})}{T_{\text{hm}}} \quad (6)$$

From Figure 6 and Supplemental Figure S4 it is apparent that, for all sites, our results are consistent with the UV/Vis melting data (Arrhenius behavior) with the notable exception of the modified base (Supplemental Figure S4) where both processes follow an Arrhenius behavior. The robustness of the CEST-based thermodynamic analysis is reflected in the exceptionally high consistency between the thermodynamic parameters when the same sites are compared.

Site-specific activation kinetic parameters for both samples are summarized in Table 1, and visualized in the insets of Figure 7. Consistently with the UV/Vis melting data, the 5fC modification destabilizes the 3-mer DNA by ~ 4 –8 kJ mol $^{-1}$ at 37°C, and decreases the melting temperature by $\sim 2^\circ\text{C}$. As a support for the two-state model, we found that the destabilization is not localized, but it involves the whole molecule. The ΔS and ΔH parameters are scattered over a wide range and seemingly independent of their absolute value or chemical type (i.e. if they are purine or pyrimidine H8 protons). The only main outlier is the formyl pyrimidine base, C6H7, whose ΔH^\ddagger and ΔS_i^\ddagger are $\sim 70\%$ larger, and ΔG_i^\ddagger and ΔS_i^\ddagger have the opposite

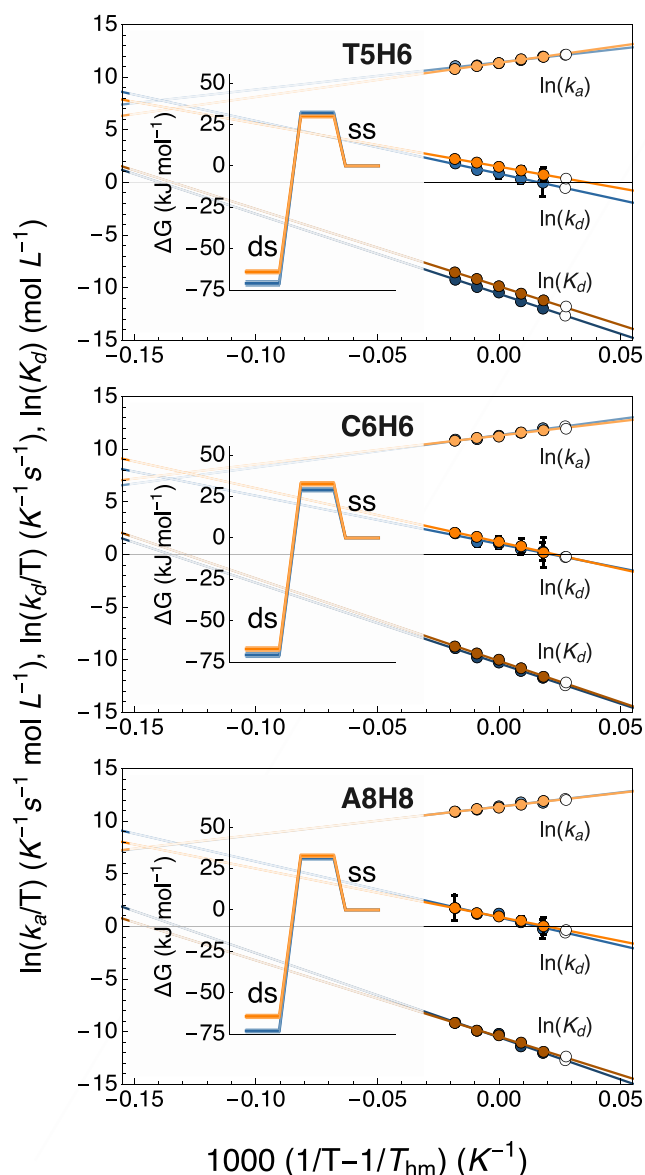


Figure 6. van't Hoff plots are displayed for three selected sites. Shades of orange and blue indicate data entries and linear fits for Mod₁₂ and Can₁₂, respectively. White data points represent back-calculated values. Fitted values for the association and dissociation rate constants (k_a and k_d) and the equilibrium constant of dissociation K_d are listed in Supplementary Tables S5 (Mod₁₂) and S6 (Can₁₂). Insets present the relevant Gibbs free energy plots at each site at 37°C.

site sign as those of the other H2, H6 and H8 protons in the molecule. To assess whether this large difference in the thermodynamics and kinetics values of C6H7 reflects a divergence from a two-state equilibrium we plotted ΔH as a function of ΔS (such a plot reflects the genuine entropy-enthalpy compensation). Since $\Delta G = \Delta H - T\Delta S$, therefore the plot of ΔH versus ΔS gives a linear with a slope of T_c which corresponds to the characteristic temperature of the measurement, and the intercept reflects the ΔG at T_c . Figure 7A–C displays the entropy-enthalpy compensation plots for both thermodynamic and kinetic processes with a fitted linear including all observable protons of both samples. Most sites

are clustered together, while C6H7, G1H8 and C2H6 protons of Mod₁₂ have distinctly different ΔH and ΔS values, however their ΔH and ΔS compensate each other leading to a ΔG , which agrees well with the overall ΔG of the system. The fitted slopes (T_c) and intercepts ($\Delta G(T_c)$) are close to the experimental T_{hm} and obtained average $\Delta G(T_{hm})$ values.

Since ΔG is the difference of two large quantities, the ΔH versus ΔS plot is ill-suited to indicate any site-specific differences in the Gibbs free energy differences. The subtle destabilization caused by the epigenetically relevant 5fC nucleotide can be better appreciated in Figure 7D where ΔG_d^\ddagger is plotted as a function of ΔG° for both samples. Here again, C6H7 seems to be an outlier with much higher activation free energies and lower equilibrium energies than the rest of the molecule. Generally, both the ΔG° and ΔG_d^\ddagger values are decreased by ~ 10 and ~ 20 kJ mol⁻¹ in Mod₁₂ with respect to the same values in Can₁₂, respectively. Otherwise, we see a linear correlation between the activation and equilibrium free energies (C6H7 was excluded from this analysis). In terms of linear Gibbs free energy relation, the slope $d\Delta G_d^\ddagger/d\Delta G^\circ$, with a value between 0 and 1, measures how much the transition state resembles the GS or ES conformations (67). In our case, the slope of the Gibbs energy changes associated with the DNA melting process is 0.83 which indicates that the transition state is structurally related to the single-stranded conformation.

DISCUSSION

The naturally occurring 5fC DNA modification is an epigenetic marker whose biological roles have not been yet precisely defined (4–7). In order to account for the impact of 5fC incorporation into dsDNA, we studied how such a modification reshapes the thermodynamics and kinetics of the process of DNA melting and annealing.

We propose that our findings can be interpreted within the framework of two separate, but complementary aspects: the role of 5fC as an independent, semi-permanent modification and its unique interactions with TDG protein within the framework of the active demethylation cycle.

5fC as a semi-permanent modification

In agreement with previous studies, we identified no indications of 5fC-induced wobble base-pairing arising from an imino-amino tautomerization processes, nor any hint pointing towards a deviation from the B-DNA structure (34–36). If there is no obvious modification on the ground-state structure of dsDNA due to 5fC incorporation then the questions arise, whether it impacts the properties of a rarely visited excited state or it alters the potential energy surface around the accessible conformations.

To address these questions, we performed a thorough, site-specific thermodynamics and kinetics study on the melting process involving two 12mer dsDNA samples. From the preliminary UV/Vis melting study, we could confirm that 5fC destabilizes the DNA structure when compared to canonical C by increasing the equilibrium constant of the dissociation, but this method was unable to clarify

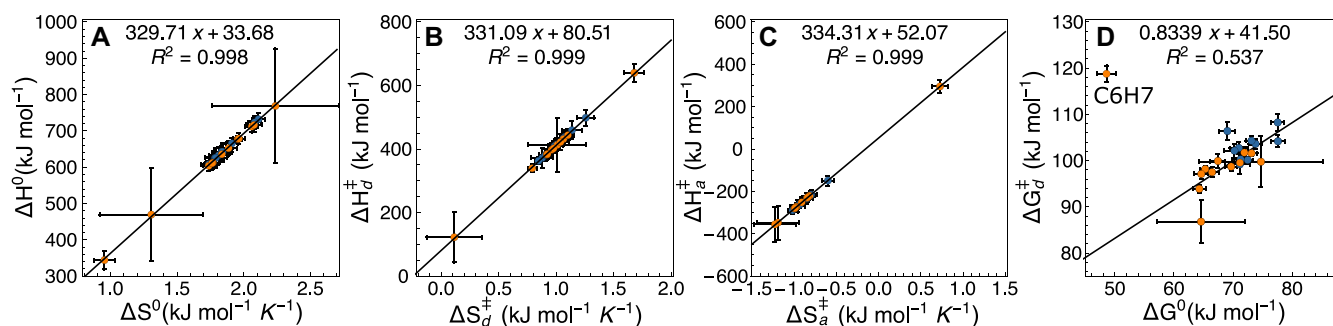


Figure 7. (A–C) Correlation between the enthalpy and entropy changes of the thermodynamic and kinetic processes for Can₁₂ (blue) and Mod₁₂ (orange). The solid lines represent a linear fit in the form of $\Delta H = T_c \Delta S + \Delta G$. (D) Correlation between the dissociation and equilibrium free energy changes of Can₁₂ (blue) and Mod₁₂ (orange) obtained at 37°C.

if this increase was a consequence of the decreased association rate, or the increased dissociation rate. Hence, we turned our attention to CEST measurements which provide site-specific information about both the kinetics (exchange rates) and the thermodynamics (populations) of simple equilibrium processes. With CEST, we could establish that the decreased stability of 5fC modified DNA was a consequence of the ~2–5-fold increase of the dissociation rate and the ~1.5–3.5-fold decrease of the association rate, leading to a ~4–9 times higher equilibrium dissociation constant.

We propose that these results are relevant for our understanding of 5fC as an independent, semi-permanent modification and can be explained considering the molecular grounds of the destabilization. Experimental and computational studies have described the electron-withdrawing (EW) properties of the formyl group on position 5 (38,39), and it is conceivable that the very same EW effects that modulate the dissociation process have a similar impact on the annealing kinetics by curtailing the extent to which the critical amount of bases needed to trigger the nucleation process, a key factor in double-stranded nucleic acid hybridization, come in contact (3).

By performing the CEST measurements in a temperature-dependent fashion we ascertained that the kinetic rates and the equilibrium constants follow an Eyring and van't Hoff relationship, respectively (cf. Figure 6 and Supplementary Figure S4). Indeed, the overall similarity of the profiles between equivalent sites of the two samples suggests that inclusion of 5fC leaves the transition states and potential energy surfaces unaltered. Consistently with previous investigations of the melting kinetics of short dsDNA sequences, the dissociation kinetics were found to be compatible with an Arrhenius process, in opposition to the annealing mechanism, which appears to progressively slow down at higher temperatures (3,65). The validity of the above observations holds for both samples, indicating that 5fC does not drastically perturb dsDNA and does not fundamentally change its hybridization and melting dynamics.

In order to understand the subtle differences arising from the incorporation of formylated cytosine in DNA, we resorted to a global analysis of the thermodynamics and kinetic parameters determined via our CEST and UV/Vis studies. The correlation of the activation free energy change

(ΔG_a^\ddagger) with the standard Gibbs free energy change (ΔG^0) assists in rationalizing the elusive 5fC-induced destabilization, which stems from the higher energy associated with the ground state of the dsDNA conformer. Indeed, values relative to Mod₁₂ tend to cluster at the lower energy side of the correlation plot, meaning that the lower energy barrier separating ds- and ssDNA is mostly, if not entirely due to the lower cohesion between the two strands. Furthermore, the impact of the destabilization generated by the centrally placed modification is not localized as previously suggested (38,68,69). Rather, it appears to be broad and all base pairs are consistently affected when compared to the canonical sample. This evidence implies that 5fC curtails the extent to which bases in dsDNA can cooperatively stabilize each other, causing a global weakening of the double stranded structure, albeit without disrupting the original B-DNA geometry. By examining the ¹H 1D spectra of Mod₁₂ and Can₁₂, it can be appreciated how the sole proton in the imino region which is shifted upon 5fC incorporation is G7H1, indicating that the hydrogen bond strength of the centrally positioned C6:G7 base pair is significantly weakened upon cytosine formylation. In sharp contrast, other base pairs, and most notably even the immediate neighbour T5:A8, are entirely unaffected in terms of H-bonding capacity. The above observation is deceptively simple when compared to the numerical results obtained via our CEST analysis. How can then such a localized effect impact global processes such as melting and reannealing? We propose that the weakened H-bond of 5fC6:G7 diminishes the extent to which the entire ensemble of bases can cooperatively stabilize each other via mechanisms of π – π stacking, cross-strands interactions etc. as previously reported (1–3). Lower H-bond propensity translates into a higher propensity for base wiggling on very fast (ns– μ s) time scales, which in turn weakens the extent to which neighboring bases stabilize each other. A schematic representation of this process is depicted in Supplementary Figures S6 and S7.

Previous research has established the notion that the rate limiting step in the process of ssDNA association dynamics is the nucleation (3,70). During the formation of a DNA double strand, the transition state structure only features a few preformed base pairs. Then, by progressively ‘zipping-up’ the remaining sites until it reaches the termini, and achieves the formation of the double strand. The nega-

tive ΔH_a^\ddagger that we observed for all but one sites corroborates this idea and indicates that the transition state towards a double-stranded structure is enthalpically stabilized by newly formed hydrogen bonds. Remarkably, the only proton which follows an opposite trend is the formyl proton of 5fC. Here, the positive ΔH_a^\ddagger suggests that the formyl group forms enthalpically favorable interactions (e.g. H-bonds) even in the single-stranded state, and these interactions should break and newly form during the hybridization process. The anomalous behavior of C6H7 with respect to other proton reporters could stem from the rotation of the formyl group along the axis of the carbonyl carbon and the covalently bonded ring carbon, which would normally occur on the nanosecond timescale. We speculate that the strong hydrogen bond between the carbonyl oxygen and the nearby amino protons slows this rotation down to the ms time scale and makes this exchange process interfere with the melting kinetics. Thus, the observed kinetics and thermodynamics of C6H7 is a mixture of both bond rotation and DNA hybridization.

The correlation between standard and activation Gibbs free energy changes has been long used to comment on the nature of the transition state (71). The slope of the fitted linear normally ranges between 0 and 1, meaning that the transition state structurally resembles the reactant or the product, respectively. Consistently with this view, we find a slope of 0.83, providing experimental support to the nucleation theory where the otherwise floppy strands are anchored only at few locations and hence resembling the single-stranded form. This finding is in contrast to the conclusion of the studies related to individual base-pair openings, which found that the transition state of that process resembles the dsDNA architecture more than the ssDNA random coil (67).

In view of the foregoing, we conclude that formylation of cytosine within its naturally occurring setting distinctly enhances dsDNA opening. It has been long established that the presence of the simplest cytosine epigenetic modification, 5mC, is linked to ‘genetic silencing’, i.e. repression of gene expression (19–21). 5mC has also been confirmed to stabilize the dsDNA conformer, both kinetically and thermodynamically (34,38,40). Our study suggests that, in sharp contrast, by enhancing dsDNA double strand opening 5fC might facilitate gene expression. The fact that formylated cytosine is abundant (relative to its presence in different types of genomes) in the context of cancerous and stem cells, where cell activity is enhanced, is perhaps more than coincidental.

5fC as a substrate

A central conundrum of epigenetics is how reader proteins recognize their substrates with such high specificity over the overwhelming background represented by canonical nucleic acids. This puzzle is especially intriguing for TDG which can differentiate between carbonyl-bearing epigenetically modified cytosines such as 5fC and 5caC versus canonical, methylated and hydroxymethylated (C, 5mC, 5hmC) cytosines. Several studies have proposed the so-called ‘pinch-push-pull’ mechanism to explain the outstanding recognition capabilities of TDG towards dsDNA mismatches (e.g.

the G–T mismatched base-pair), 5fC and 5caC (although in a highly pH-dependent manner) (72–75). According to the consensus enzymatic model for TDG substrate recognition, base extrusion is proposed to be initiated by residue Arg275, which interferes with Watson-Crick base-pairing geometry and triggers the process of substrate flipping into the active site of TDG. An array of subsequent recognition and binding processes precedes the enzymatic event of the glycosidic bond cleavage. Although our work does not involve studies of DNA-protein interactions, and thus is ill-suited to comment on the whole enzymatic process, we propose that our results fit nicely within the above mentioned ‘pinch-push-pull’ mechanism. Indeed, we suggest that the lower activation barrier at 37°C that 5fC-bearing DNA encounters in the course of a melting process facilitates the base extrusion enacted by TDG. Upon interrogation of Arg275, the weakened (by the inclusion of 5fC) Watson-Crick base pair geometries facilitate base flipping into the active site of the enzyme. In this context, canonical bases (or other epigenetic modifications lacking the carbonyl moiety which are not recognized by TDG, such as 5mC and 5hmC) are then released and resume their Watson-Crick hydrogen bonding. Otherwise, 5fC is recognized and retained in the active site, allowing TDG to perform base-excision.

CONCLUSION

In this work, we investigated the effect of the inclusion of 5fC within its naturally occurring settings and conditions in the context of a double-stranded DNA 12mer, focusing our efforts on the slow time-scale dynamics close to the melting temperature.

We achieved a quantitative assessment of the degree of destabilization degree by 5fC via ^1H CEST techniques, which yielded the populations of both the single and double stranded DNA conformers, their chemical shifts, and the overall exchange kinetic rates k_{ex} for each temperature. Further decomposition of the site-specific k_{ex} into k_d and k_a allowed for the determination of kinetic rates for both the melting and annealing processes. We find that both k_d and k_a are consistently affected by the presence of 5fC and their combined impact on the dissociation equilibrium constant near the melting temperature can be almost as large as an order of magnitude. To gain further insights into the energetics of double helix to single helix transition, we developed a methodological framework for the extraction of thermodynamic (ΔG° , ΔH° , ΔS° , T_m) and kinetic (ΔG^\ddagger , ΔH^\ddagger , ΔS^\ddagger) parameters from the CEST derived k_{ex} and population values. Our findings indicate that inclusion of 5fC generates a tenuous degree of destabilization in the duplex state without affecting the transition state architecture, which remains structurally indistinguishable from its canonical counterpart. Said destabilization affects the conformational dsDNA-ssDNA conformational equilibrium. We speculate the identified features induced by 5fC onto dsDNA might account for the both its role as an independent, semi-permanent epigenetic marker and the selective recognition and activity of the TDG protein towards sites bearing the modification.

In this study, we confined our analysis to slow ms time-scale motions, focusing on the site-specific quantification of

5fC-related energetics. The noticeable and localized weakening of the 5fC:G base pair in contrast to its canonical C:G counterpart implies that the origin of the induced destabilization effects has its roots in a faster time-scale motion, which has as of yet gone undetected. Further research will be needed to unravel and characterize possible excited states on faster time scales that might be impacting in a localized fashion the dsDNA flexibility at (or near) the 5fC:G base pair.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Anthony Crisp for insightful discussion and help in the measurements of UV/Vis melting profiles.

FUNDING

Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) [SFB 1309-325871075, in part]; EU-ITN LightDyNAmics [765266]; ERC-AG EpiR [741912]; Center for NanoScience, the Excellence Clusters CIPSM; Fonds der Chemischen Industrie. Funding for open access charge: The SFB1309 framework (<https://www.sfb1309.de/about-sfb-1309/>).

Conflict of interest statement. None declared.

REFERENCES

- Poater, J., Swart, M. and Bickelhaupt, F.M. (2014) B-DNA structure and stability: the role of hydrogen bonding, π - π stacking interactions, twist-angle, and solvation. *Org. Biomol. Chem.*, **12**, 4691–4700.
- Hobza, P. and Šponer, J. (1999) Structure, energetics, and dynamics of the nucleic acid base pairs: nonempirical ab initio calculations. *Chem. Rev.*, **99**, 3247–3276.
- Ouldrige, T.E., Šulc, P., Romano, F., Doye, J.P.K. and Louis, A.A. (2013) DNA hybridization kinetics: zippering, internal displacement and sequence dependence. *Nucleic Acids Res.*, **41**, 8886–8895.
- Bilyard, M.K., Becker, S. and Balasubramanian, S. (2020) Natural, modified DNA bases. *Curr. Opin. Chem. Biol.*, **57**, 1–7.
- Carell, T., Kurz, M.Q., Müller, M., Rossa, M. and Spada, F. (2018) Non-canonical bases in the genome: the regulatory information layer in DNA. *Angew. Chem. Int. Ed.*, **57**, 4296–4312.
- Chen, Y., Hong, T., Wang, S., Mo, J., Tian, T. and Zhou, X. (2017) Epigenetic modification of nucleic acids: from basic studies to medical applications. *Chem. Soc. Rev.*, **46**, 2844–2872.
- Hardwick, J.S., Lane, A.N. and Brown, T. (2018) Epigenetic modifications of cytosine: biophysical properties, regulation, and function in mammalian DNA. *BioEssays*, **40**, doi:10.1002/bies.201700199.
- Sasaki, H. and Matsui, Y. (2008) Epigenetic events in mammalian germ-cell development: reprogramming and beyond. *Nat. Rev. Genet.*, **9**, 129–140.
- Yamaguchi, S., Hong, K., Liu, R., Inoue, A., Shen, L., Zhang, K. and Zhang, Y. (2013) Dynamics of 5-methylcytosine and 5-hydroxymethylcytosine during germ cell reprogramming. *Cell Res.*, **23**, 329–339.
- Neri, F., Incarnato, D., Krepelova, A., Rapelli, S., Anselmi, F., Parlato, C., Medana, C., Dal Bello, F. and Oliviero, S. (2018) Single-base resolution analysis of 5-formyl and 5-carboxyl cytosine reveals promoter DNA methylation dynamics. *Cell Rep.*, **10**, 674–683.
- Basanta-Sanchez, M., Wang, R., Liu, Z., Ye, X., Li, M., Shi, X., Agris, P.F., Zhou, Y., Huang, Y. and Sheng, J. (2017) TET1-mediated oxidation of 5-formylcytosine (5fC) to 5-carboxycytosine (5caC) in RNA. *ChemBioChem*, **18**, 72–76.
- Van Haute, L., Powell, C.A. and Minczuk, M. (2017) Dealing with an unconventional genetic code in mitochondria: the biogenesis and pathogenic defects of the 5-formylcytosine modification in mitochondrial tRNA^(Met). *Biomolecules*, **7**, 24.
- Dawson, M.A. and Kouzarides, T. (2012) Cancer epigenetics: from mechanism to therapy. *Cell*, **150**, 12–27.
- Cavalli, G. and Heard, E. (2019) Advances in epigenetics link genetics to the environment and disease. *Nature*, **517**, 489–499.
- Baylin, S. and Jones, P.A. (2011) A decade of exploring the cancer epigenome – biological and translational implications. *Nat. Rev. Cancer*, **11**, 726–734.
- Wu, H. and Zhang, Y. (2011) Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation. *Genes Dev.*, **25**, 2436–2452.
- Storebjerg, T.M., Strand, S.H., Høyer, S., Lynnerup, A., Borre, M., Ørntoft, T.F. and Sørensen, K.D. (2018) Dysregulation and prognostic potential of 5-methylcytosine (5mC), 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxycytosine (5caC) levels in prostate cancer. *Clin. Epigenet.*, **10**, 105.
- Derreumaux, S., Chaoui, M., Tevianian, G. and Femandjian, S. (2001) Impact of CpG methylation on structure, dynamics and solvation of cAMP DNA responsive element. *Nucleic Acids Res.*, **29**, 2314–2326.
- Bird, A. (2002) DNA methylation patterns and epigenetic memory. *Genes Dev.*, **16**, 6–21.
- Moarii, M., Boeva, V., Vert, J. and Reyal, F. (2015) Changes in correlation between promoter methylation and gene expression in cancer. *BMC Genomics*, **16**, 873.
- Jones, P.A. (2012) Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.*, **13**, 484–492.
- Smith, Z.D. and Meissner, A. (2013) DNA methylation: roles in mammalian development. *Nat. Rev. Genet.*, **14**, 204–220.
- Koivunen, P. and Laukka, T. (2018) The TET enzymes. *Cell. Mol. Life Sci.*, **75**, 1339–1348.
- Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C. and Zhang, Y. (2011) Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxycytosine. *Science*, **333**, 1300–1304.
- Hu, L., Li, Z., Cheng, J., Rao, Q., Gong, W., Liu, M., Shi, Y.G., Zhu, J., Wang, P. and Xu, Y. (2013) Crystal structure of TET2-DNA complex: insight into TET-mediated 5mC oxidation. *Cell*, **155**, 1545–1555.
- Pfaffeneder, T., Hackner, B., Truß, M., Münzel, M., Müller, M., Deiml, C.A., Hagemeyer, C. and Carell, T. (2011) The discovery of 5-formylcytosine in embryonic stem cell DNA. *Angew. Chem. Int. Ed.*, **50**, 7008–7012.
- Su, M., Kirchner, A., Stazzoni, S., Müller, M., Wagner, M., Schröder, A. and Carell, T. (2016) 5-Formylcytosine could be a semipermanent base in specific genome sites. *Angew. Chem. Int. Ed.*, **55**, 11797–11800.
- Bachman, M., Uribe-Lewis, S., Yang, X., Burgess, H.E., Iurlaro, M., Reik, W., Murrell, A. and Balasubramanian, S. (2015) 5-Formylcytosine can be a stable DNA modification in mammals. *Nat. Chem. Biol.*, **11**, 555–557.
- Zhang, Y. and Zhou, C. (2019) Formation and biological consequences of 5-Formylcytosine in genomic. *DNA Repair*, **81**, 102649.
- Pfeifer, G.P., Szabó, P.E. and Song, J. (2020) Protein interactions at oxidized 5-methylcytosine bases. *J. Mol. Biol.*, **432**, 1718–1730.
- Rausch, C., Hastert, F.D. and Cardoso, M.C. (2019) DNA modification readers and writers and their interplay. *J. Mol. Biol.*, **432**, 1731–1746.
- Wu, X. and Zhang, Y. (2017) TET-mediated active DNA demethylation: mechanism, function and beyond. *Nat. Rev. Genet.*, **18**, 517–534.
- Raiber, E.A., Murat, P., Chirgadze, D.Y., Beraldi, D., Luisi, B.F. and Balasubramanian, S. (2015) 5-formylcytosine alters the structure of the DNA double helix. *Nat. Struct. Mol. Biol.*, **22**, 44–49.
- Szulik, M.W., Pallan, P.S., Nocek, B., Voehler, M., Banerjee, S., Brooks, S., Joachimiak, A., Egli, M., Eichman, B.F. and Stone, M.P. (2015) Differential stabilities and sequence-dependent base pair opening dynamics of Watson Crick base pairs with 5-hydroxymethylcytosine, 5-formylcytosine, or 5-carboxycytosine. *Biochemistry*, **134**, 8148–8161.
- Hardwick, J.S., Ptchelkine, D., El-Sagheer, A.H., Tear, I., Singleton, D., Phillips, S.E.V., Lane, A.N. and Brown, T. (2017) 5-Formylcytosine does not change the global structure of DNA. *Nat. Struct. Mol. Biol.*, **24**, 544–552.

36. Fu, T., Liu, L., Yang, Q., Wang, Y., Xu, P., Zhang, L., Liu, S., Dai, Q., Ji, Q., Xu, G. *et al.* (2019) Thymine DNA glycosylase recognizes the geometry alteration of minor grooves induced by 5-formylcytosine and 5-carboxylcytosine. *Chem. Sci.*, **10**, 7407–7417.
37. Ngo, T.T.M., Yoo, J., Dai, Q., Zhang, Q., He, C., Aksimentiev, A. and Ha, T. (2016) Effects of cytosine modifications on DNA flexibility and nucleosome mechanical stability. *Nat. Commun.*, **7**, 10813.
38. Dai, Q., Sanstead, P.J., Peng, C.S., Han, D. and He, C. (2016) Weakened N3 hydrogen bonding by 5-formylcytosine and 5-carboxylcytosine reduces their base-pairing stability. *ACS Chem. Biol.*, **11**, 470–477.
39. Nunes, F.B., Bettega, M.H.F. and Almeida Sanchez, S. (2017) Substituents' effect in electron attachment to epigenetic modifications of cytosine. *J. Chem. Phys.*, **146**, 244314.
40. Sanstead, P.J., Ashwood, B., Dai, Q., He, C. and Tokmakoff, A. (2020) Oxidized derivatives of 5-methylcytosine alter the stability and dehybridization dynamics of duplex DNA. *J. Phys. Chem. B*, **124**, 1160–1174.
41. Mulder, F.A.A., Mittermaier, A., Hon, B., Dahlquist, F.W. and Kay, L.E. (2001) Studying excited states of proteins by NMR. *Nat. Struct. Mol. Biol.*, **11**, 932–935.
42. Nikolova, E.N., Kim, E., Wise, A.A., Brien, P.J.O., Andricioaei, I. and Al-Hashimi, H.M. (2011) Transient Hoogsteen base pairs in canonical duplex DNA. *Nature*, **470**, 498–502.
43. Tollinger, M., Skrynnikov, N.R., Mulder, F.A.A., Forman-Kay, J.D. and Kay, L.E. (2001) Slow dynamics in folded and unfolded states of an SH3 domain. *J. Am. Chem. Soc.*, **123**, 11341–11352.
44. Marušič, M., Schlagntweit, J. and Petzold, K. (2019) RNA dynamics by NMR spectroscopy. *ChemBioChem*, **20**, 2685–2710.
45. Zhou, H., Kimsey, I.J., Nikolova, E.N., Sathyamoorthy, B., Grazioli, G., McSally, J., Bai, T., Wunderlich, C.H., Kreutz, C., Andricioaei, I. *et al.* (2016) m¹A and m¹G disrupt A-RNA structure through the intrinsic instability of Hoogsteen base pairs. *Nat. Struct. Mol. Biol.*, **23**, 803–810.
46. Juen, M.A., Wunderlich, C.H., Nußbaumer, F., Tollinger, M., Kontaxis, G., Konrat, R., Hansen, D.F. and Kreutz, C. (2016) Excited states of nucleic acids probed by proton relaxation dispersion NMR spectroscopy. *Angew. Chem. Int. Ed.*, **55**, 12008–12012.
47. Nikolova, E.N., Gottardo, F.L. and Al-Hashimi, H.M. (2012) Probing transient Hoogsteen hydrogen bonds in canonical duplex DNA using NMR relaxation dispersion and single-atom substitution. *J. Am. Chem. Soc.*, **8**, 3667–3670.
48. Dethoff, E.A., Petzold, K., Chugh, J., Casiano-Negroni, A. and Al-Hashimi, H.M. (2012) Visualizing transient low-populated structures of RNA. *Nature*, **491**, 724–728.
49. Vallurupalli, P., Bouvignies, G. and Kay, L.E. (2012) Studying 'Invisible' excited protein states in slow exchange with a major state conformation. *J. Am. Chem. Soc.*, **134**, 8148–8161.
50. Vinogradov, E., Sherry, A.D. and Lenkinski, R.E. (2013) CEST: from basic principles to applications, challenges and opportunities. *J. Magn. Reson.*, **229**, 155–172.
51. Palmer III, A.G. and Koss, H. (2019) In: *Methods in Enzymology*. Springer, Berlin, pp. 177–227.
52. Zhao, B., Baisden, J.T. and Zhang, Q. (2020) Probing excited conformational states of nucleic acids by nitrogen CEST NMR spectroscopy. *J. Magn. Reson.*, **310**, 106642.
53. Delaforge, E., Milles, S., Bouvignies, G., Bouvier, D., Boivin, S., Salvi, N., Maurin, D., Martel, A., Round, A., Lemke, E.A. *et al.* (2015) Large-scale conformational dynamics control H5N1 influenza polymerase PB2 binding to Importin α . *J. Am. Chem. Soc.*, **137**, 15122–15134.
54. Shi, H., Liu, B., Nussbaumer, F., Rangadurai, A., Kreutz, C. and Al-Hashimi, H.M. (2019) NMR chemical exchange measurements reveal that N⁶-methyladenosine slows RNA annealing. *J. Am. Chem. Soc.*, **51**, 19988–19993.
55. Schlagntweit, J., Steiner, E., Karlsson, H. and Petzold, K. (2018) Efficient detection of structure and dynamics in unlabeled RNAs: the SELOPE approach. *Chem. Eur. J.*, **24**, 6067–6070.
56. Schröder, A.S., Steinbacher, J., Steigenberger, B., Gnerlich, F.A., Schiesser, S., Pfaffeneder, T. and Carell, T. (2014) Synthesis of a DNA promoter segment containing all four epigenetic nucleosides: 5-methyl-, 5-hydroxymethyl-, 5-formyl-, and 5-carboxy-2'-deoxycytidine. *Angew. Chem. Int. Ed.*, **53**, 315–318.
57. Lee, W., Tonelli, M. and Markley, J.L. (2015) NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics*, **31**, 1325–1327.
58. Kim, S., Lin, L. and Reid, B.R. (1992) Determination of nucleic acid backbone conformation by proton NMR. *Biochemistry*, **31**, 782–791.
59. Hilal, S.H., Bornander, L.L. and Carreira, L.A. (2005) Hydration equilibrium constants of aldehydes, ketones and quinoxalines. *QSAR Comb. Sci.*, **24**, 631–638.
60. Johnson, R.P., Fleming, A.M., Perera, R.T., Burrows, C.J. and White, H.S. (2017) Dynamics of a DNA mismatch site held in confinement discriminate epigenetic modifications of cytosine. *J. Am. Chem. Soc.*, **139**, 2750–2756.
61. Mergny, J. and Lacroix, L. (2003) Analysis of thermal melting curves. *Oligonucleotides*, **13**, 515–537.
62. Dragan, A., Privalov, P. and Crane-Robinson, C. (2019) Thermodynamics of DNA: heat capacity changes on duplex unfolding. *Eur. Biophys. J.*, **48**, 773–779.
63. Lam, S.L., Ip, L.N., Cui, X. and Ho, C.N. (2002) Random coil proton chemical shifts of deoxyribonucleic acids. *J. Biomol. NMR*, **24**, 329–337.
64. Lam, S.L. (2007) DSHIFT: a web server for predicting DNA chemical shifts. *Nucleic Acid Res.*, **35**, W713–W717.
65. Menssen, R.J. and Tokmakoff, A. (2019) Length-dependent melting kinetics of short DNA oligonucleotides using temperature-jump IR spectroscopy. *J. Phys. Chem. B*, **123**, 756–767.
66. Nöltig, B. (2006) In: *Protein Folding Kinetics, Biophysical Methods*. Springer, Berlin.
67. Coman, D. and Russu, I.M. (2005) A nuclear magnetic resonance investigation of the energetics of basepair opening pathways in DNA. *Biophys. J.*, **89**, 3285–3292.
68. Hashimoto, H., Hong, S., Bhagwat, A.S., Zhang, X. and Cheng, X. (2012) Excision of 5-hydroxymethyluracil and 5-carboxylcytosine by the thymine DNA glycosylase domain: its structural basis and implications for active DNA demethylation. *Nucleic Acid Res.*, **40**, 10203–10214.
69. Maiti, A. and Drohat, A.C. (2011) Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine. *J. Biol. Chem.*, **286**, 35334–35338.
70. Sicard, F., Destainville, N. and Manghi, M. (2015) DNA denaturation bubbles: free-energy landscape and nucleation/closure rates. *J. Chem. Phys.*, **142**, 034903.
71. Leffler, J.E. (1953) Parameters for the description of transition states. *Science*, **117**, 340–341.
72. Coey, C.T., Malik, S.S., Pidugu, L.S., Varney, K.M., Pozharski, E. and Drohat, A.C. (2016) Structural basis of damage recognition by thymine DNA glycosylase: key roles for N-terminal residues. *Nucleic Acids Res.*, **44**, 10248–10258.
73. Maiti, A., Morgan, M.T., Pozharski, E. and Drohat, A.C. (2008) Crystal structure of human thymine DNA glycosylase bound to DNA elucidates sequence-specific mismatch recognition. *Proc. Natl. Acad. Sci. U.S.A.*, **26**, 8890–8895.
74. Dodd, T., Yan, C., Kossmann, B.R., Martin, K. and Ivanov, I. (2018) Uncovering universal rules governing the selectivity of the archetypal DNA glycosylase TDG. *Proc. Natl. Acad. Sci. U.S.A.*, **115**, 5974–5979.
75. Drohat, A.C. and Coey, C.T. (2016) Role of base excision 'repair' enzymes in erasing epigenetic marks from DNA. *Chem. Rev.*, **116**, 12711–12729.

5.3. Synthesis of galactosyl-queuosine and distribution of hypermodified Q-nucleosides in mouse tissues

Peter Thumbs[#], Timm T. Ensfelder[#], Markus Hillmeier[#], Mirko Wagner[#], Matthias Heiss, Constanze Scheel, Alexander Schön, Markus Müller, Stylianos Michalakis, Stefanie Kellner and Thomas Carell*

([#] geteilte Erstautorenschaft)

Prolog

Die Strukturaufklärung bisher nicht exakt charakterisierter RNA-Modifikationen ist von enormer Wichtigkeit für das schlussendliche Verständnis des Auftretens und der Funktionen der jeweiligen Modifikationen. Es ist bekannt, dass die RNA-Hypermodifikation Queuosin in bestimmten tRNAs (tRNA^{Tyr} und tRNA^{Asp}) von Eukaryoten durch Glykosylierung mit Mannose oder Galaktose sogar noch weiter modifiziert werden kann. Allerdings sind die exakten regio- und stereochemischen Strukturen dieser glykosylierten Nukleoside nicht weiter bekannt.

Im Rahmen dieser Arbeit wurde die Totalsynthese von β -Galaktosylqueuosin (β -GalQ) entwickelt. Durch Koinjektionsexperimente des synthetischen Standards mit dem aus tRNA isolierten Naturstoff unter Verwendung hochaufgelöster UHPLC-Massenspektrometrie konnte die Struktur des natürlich vorkommenden Moleküls aufgeklärt werden. Darüber hinaus wurden die Konzentrationen der Q-Nukleoside durch Untersuchung von Gewebeproben verschieden alter Mäuse organ- und altersspezifisch quantifiziert.

Autorenbeitrag

Peter Thumbs entwickelte die Synthese von β -GalQ und führte diese im Rahmen seiner Dissertation bis zum Jahr 2013 durch.^[104] Ich führte auf Grundlage der Vorschriften und von synthetisierten Vorstufen von *Dr. Peter Thumbs* eine Resynthese von β -GalQ durch, das nun für die Koinjektionsexperimente verwendet werden konnte, da das damals synthetisierte Material nicht mehr vorhanden war. Viele Synthesen von Vorstufen wurden auch von *Markus Hillmeier* durchgeführt, dessen Masterarbeit ich im Rahmen dieses Projekts betreute.

Lizenz

Kopie der *Open Access* Publikation auf Basis der *Creative Commons Attribution-NonCommercial License (CC BY-NC 4.0)*, die den nicht-kommerziellen Abdruck des Artikels erlaubt.

Artikel einsehbar unter:

<https://onlinelibrary.wiley.com/doi/full/10.1002/anie.202002295>

Lizenz einsehbar unter:

<https://creativecommons.org/licenses/by-nc/4.0/>



Hypermodified Nucleosides Hot Paper

How to cite: *Angew. Chem. Int. Ed.* **2020**, 59, 12352–12356

International Edition: doi.org/10.1002/anie.202002295

German Edition: doi.org/10.1002/ange.202002295

Synthesis of Galactosyl-Queuosine and Distribution of Hypermodified Q-Nucleosides in Mouse Tissues

Peter Thumbs⁺, Timm T. Ensfielder⁺, Markus Hillmeier⁺, Mirko Wagner⁺, Matthias Heiss, Constanze Scheel, Alexander Schön, Markus Müller, Stylianos Michalakakis, Stefanie Kellner, and Thomas Carell*

In memory of Rolf Huisgen

Abstract: Queuosine (Q) is a hypermodified RNA nucleoside that is found in tRNA^{His}, tRNA^{Asn}, tRNA^{Tyr}, and tRNA^{Asp}. It is located at the wobble position of the tRNA anticodon loop, where it can interact with U as well as C bases located at the respective position of the corresponding mRNA codons. In tRNA^{Tyr} and tRNA^{Asp} of higher eukaryotes, including humans, the Q base is for yet unknown reasons further modified by the addition of a galactose and a mannose sugar, respectively. The reason for this additional modification, and how the sugar modification is orchestrated with Q formation and insertion, is unknown. Here, we report a total synthesis of the hypermodified nucleoside galactosyl-queuosine (galQ). The availability of the compound enabled us to study the absolute levels of the Q-family nucleosides in six different organs of newborn and adult mice, and also in human cytosolic tRNA. Our synthesis now paves the way to a more detailed analysis of the biological function of the Q-nucleoside family.

In all three domains of life, RNA contains next to the canonical bases (A, C, G, and U) a large variety of modified nucleosides.^[1] Most of these are chemically simple derivatives

of the canonical nucleosides. They often carry methylations at various positions of the heterocycle or the sugar, but others are heavily modified, involving multistep biosynthesis pathways. Queuosine **1** (Q) is one of the most complex of these so-called hypermodified nucleosides. (Figure 1). It is found in a large number of different species and also present in the cytosolic and mitochondrial tRNA^{Tyr}, tRNA^{Asp}, tRNA^{His}, and tRNA^{Asn} of humans.^[2–8] Interestingly, in the human cytosolic tRNA^{Tyr} and tRNA^{Asp}, Q is further modified with galactose (galQ) and mannose (manQ), respectively.^[9,10] In these tRNAs, the sugar is proposed to be attached to the homoallylic hydroxyl group of the cyclopentene ring system that is linked to the 7-deazaheterocycle.^[11] While the chemical synthesis of Q has been achieved,^[12–14] no reports exist about the preparation of its sugar-modified derivatives galQ **2** and manQ **3**, which has hampered investigations of their biological role. Accordingly, the exact function of galQ and manQ as part of the human cytosolic tRNA^{Tyr} and tRNA^{Asp} is unknown. In addition, we do not know to which extent the corresponding tRNAs are modified with different Q-family nucleosides, and how the G/Q-exchange process and the sugar derivatization is orchestrated. Furthermore, quantitative data about Q-modification levels in different organs is also lacking.

To address these questions, we performed the first total synthesis of galactosyl-queuosine **2**. This allowed us to confirm its proposed structure and to report the absolute levels of all Q-family members in different tissues of newborn and adult mice. Finally, we were able to measure to which extent human cytosolic tRNAs are modified with the three Q-family nucleosides.

[*] Dr. P. Thumbs,^[†] M. Sc. T. T. Ensfielder,^[†] M. Sc. M. Hillmeier,^[†] Dr. M. Wagner,^[†] M. Sc. M. Heiss, M. Sc. A. Schön, Dr. M. Müller, Dr. S. Kellner, Prof. Dr. T. Carell
Department of Chemistry
Ludwig-Maximilians-Universität München
Butenandtstr. 5–13, 81377 Munich (Germany)
E-mail: Thomas.Carell@lmu.de
Homepage: <https://www.carellgroup.de>
M. Sc. C. Scheel, Prof. Dr. S. Michalakakis
Department of Pharmacy
Ludwig-Maximilians-Universität München
Butenandtstr. 5–13, 81377 Munich (Germany)
Prof. Dr. S. Michalakakis
Department of Ophthalmology
Ludwig-Maximilians-Universität München
Mathildenstr. 8, 80336 Munich (Germany)

[†] These authors contributed equally to this work.

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under: <https://doi.org/10.1002/anie.202002295>.

© 2020 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

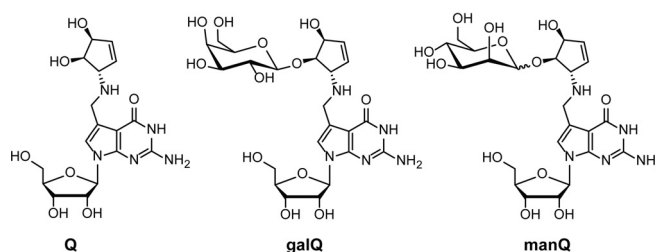


Figure 1. Depiction of the hypermodified RNA nucleoside queuosine (**1**, Q) and of the galactosylated and mannosylated Q derivatives galQ (**2**) and manQ (**3**) present in human cytosolic tRNA^{Tyr} and tRNA^{Asp}, respectively.

Galactosyl-Q **2** was constructed from three appropriately protected parts (Figure 2): The 7-formyl-7-deazaguanosine **6** was prepared, as reported by us, with Bz-protected hydroxyl groups at the ribose, and a pivaloylate protection group at the 2-amino residue.^[14] The galactose sugar was introduced as a TBS- and 2-chloroisobutyryl-protected trichloroacetimidate **4**, and the cyclopentene unit **5** was used with Fmoc-protected allyl amine and a TBS-protected allylic alcohol. We choose the 2-chlorobutyryl protecting group for the sugar-donor **4** because of its bulkiness in order to avoid unwanted orthoester formation as the main product of the glycosylation reaction, a strategy reported by Szpilman et al.^[15]

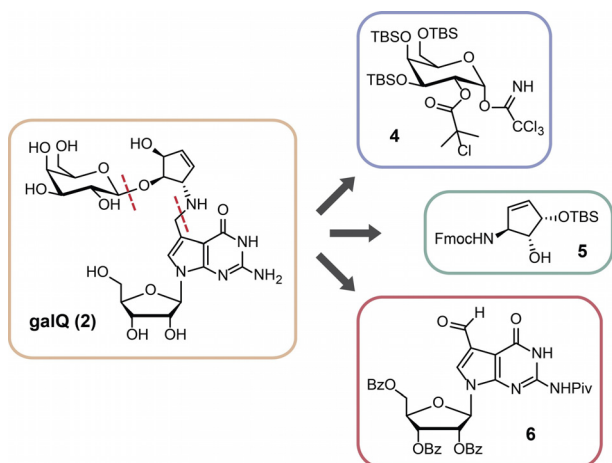
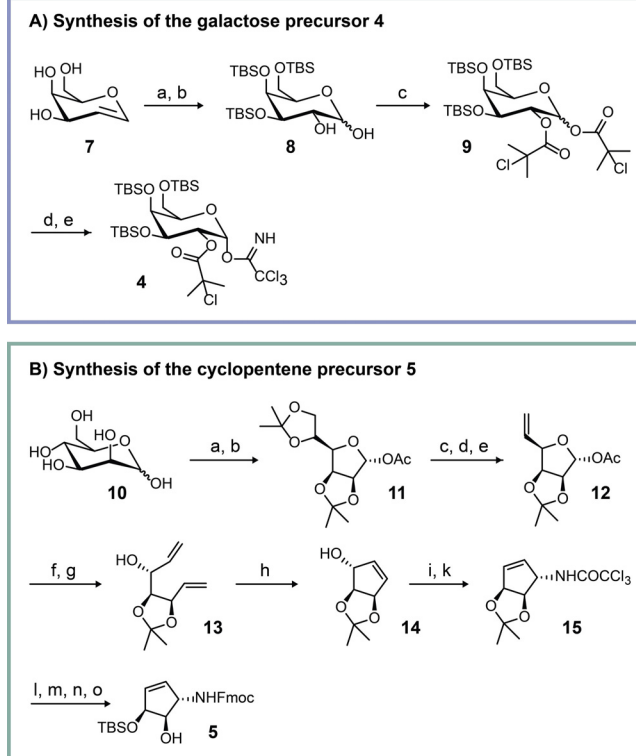


Figure 2. Retrosynthetic analysis for galQ **2**, showing the three key precursors **4**, **5**, and **6**.

The galactosyl-donor **4** itself was prepared from D-galactal **7**, which was first TBS-protected (Scheme 1 A).^[16] *cis*-Dihydroxylation of the double bond from the sterically less shielded side furnished compound **8**.^[17] This step was followed by protection of the two newly introduced hydroxyl groups with 2-chloroisobutyric acid to give the galactose-donor precursor **9**. Deprotection of the anomeric hydroxyl group with hydrazine provided the galactose precursor with a free anomeric hydroxyl group which was subsequently converted into the trichloroacetimidate donor **4** using a standard procedure.

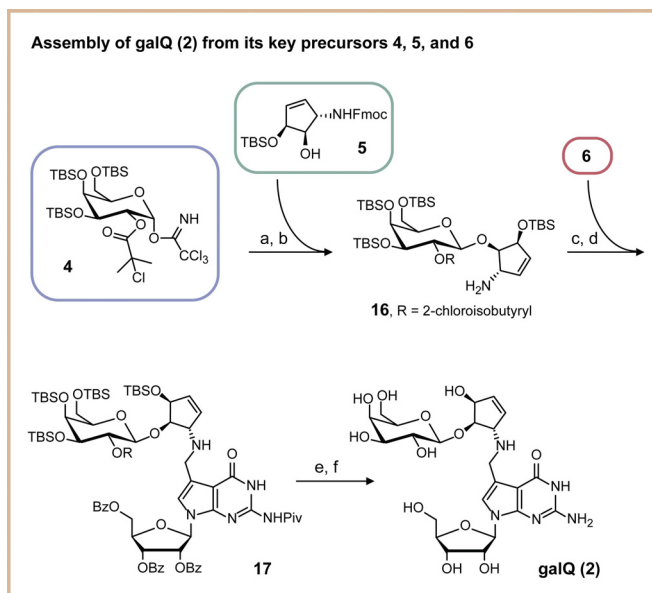
Scheme 1 B shows the synthesis of the protected 5(*S*)-amino-3(*S*),4(*R*)-dihydroxycyclopent-1-ene **5**. The starting point was mannose **10**,^[18] which was converted as reported into the double-acetonide-protected mannofuranoside **11** with an acetyl-protected anomeric center in two steps. Selective cleavage of the acetonide protecting group at the primary hydroxyl group, followed by an orthoester-based elimination, allowed introduction of a terminal double bond (**12**). Anomeric deprotection, followed by a Wittig reaction, provided the precursor **13** for the ring-closing metathesis reaction. The free hydroxyl group in **14** was then the starting point for an Overman rearrangement, providing the amine protected as the trichloroacetamide-protected amine **15**.^[19] Cleavage of this protecting group with NaOH was followed by Fmoc protection of the free amine using a standard proce-



Scheme 1. Synthesis of the key precursors **4** and **5**. A) Synthesis of the galactose precursor **4**: a) TBSCl, imidazole, DMF, 55 °C, 2 d; b) K₂OsO₄·2 H₂O, NMO, THF, *t*-BuOH, H₂O, rt, 4 h; c) 2-chloroisobutyric acid, DIC, DMAP, 0 °C, 30 min → rt, 2 h; d) N₂H₄·AcOH, DMF, −40 °C to rt, 3 h; e) Cl₃CCN, Cs₂CO₃, DCM, rt, 4 h. B) Synthesis of the cyclopentene precursor **5**: a) 2,2-dimethoxypropane, acetone, *p*-TsOH, rt, 1 h; b) Ac₂O, pyridine, 0 °C → rt, 18 h; c) aq. AcOH (66%), 55 °C, 4 h; d) triethylorthoformate, 100 °C, 30 min; e) Ac₂O, 130 °C, 5 h; f) *t*-BuOK, MeOH, 20 min; g) NaH, DMSO, Ph₃PMeBr, THF, rt → 68 °C, 2 h; h) Grubbs(I) catalyst, DCM, rt, 26 h; i) Cl₃CCN, DBU, DCM, rt, 20 min; k) *o*-xylene, 150 °C, 5 h; l) NaOH, MeOH, rt, overnight; m) Fmoc-OSu, NaHCO₃, H₂O, 1,4-dioxane; n) AcOH, H₂O, EtOAc, 50 °C, 24 h; o) TBSTf, DMF, −55 °C, 15 min.

cedure. We finally opened the acetal and protected the allylic hydroxyl group selectively with TBS-OTf in DMF at −55 °C. In this reaction, the temperature is particularly important. When the reaction was performed at higher temperatures and with prolonged reaction times, we noted selective protection of the homoallylic position.

The assembly of the galQ nucleoside **2** from the precursors **4–6** is shown in Scheme 2. We first galactosylated the cyclopentene derivative **5**. This sterically demanding step was successfully achieved by activation of the trichloroacetimidate with 2-chloro-6-methylpyridinium triflate in dichloromethane at room temperature.^[15,20] We achieved selective formation of the β-configured galactoside due to the neighboring-group effect. Subsequent cleavage of the Fmoc protection group gave product **16**, which was followed by a two-step reductive amination. First, the imine was formed in benzene, subsequently followed by reduction of the imine with NaBH₄ in methanol to afford protected galQ **17**. In a two-step deprotection protocol, we first removed the TBS groups with HF·NEt₃, followed by cleavage of ester-type



Scheme 2. Depiction of the galQ **2** assembly from the three key precursors **4**, **5**, and **6**. a) 2-Chloro-6-methylpyridinium triflate, DCM, rt, 2 h; b) DBU, MeCN, rt, 1.25 h; c) benzene, rt, 5 h; d) NaBH₄, MeOH, 0°C, 1 h; e) HF·NEt₃, DCM, rt, 4 d; f) NaOMe, MeOH, rt, 2 d.

protecting groups under Zemplén conditions. For the cleavage of the pivaloyl amide protecting group, we needed to use 0.5 M NaOMe in methanol. This strategy provided the target compound **2** with an overall yield of 0.5 % in 20 linear steps from the mannose starting molecule for the cyclopentene unit. The synthesis provided a sufficient amount of material for all further investigations.

We next investigated whether our synthetic β -homoallylic galQ **2** is identical with the natural product, because analytical data available for galQ was very limited.^[11] For this experiment, we isolated total RNA from mouse liver and performed an enzymatic digestion of the isolated RNA to the nucleoside level. This nucleoside mixture was analyzed by HPLC-MS. Indeed, under our HPLC conditions, we detected two signals with the appropriate m/z value for galQ and manQ in the extracted ion chromatogram with a retention time of around 32 and 35 min. No other peaks were present in the same m/z range. We next co-injected our synthetic β -homoallylic galQ **2**, which led to a marked increase of the second signal with a retention time of about 35 min (Figure 3). This result unambiguously showed that our synthetic compound galQ **2** and the co-eluting natural compound with the same mass are identical. Therefore, this natural compound is indeed a β -galactosylated Q derivative. Taken together, our experiment confirms the proposed chemical structure of galQ, in which the bond between the homoallylic hydroxyl group of queuosine and galactose is in β -configuration.

Having identified the HPLC retention time of galQ and therefore also of manQ, we finally were able to determine the absolute levels of galQ, manQ, and Q in different tissues of newborn (postnatal day 1; pd1) and adult mice (postnatal month 3; pm3). For an initial broad study, we measured the respective nucleoside levels in cortex, cerebellum, liver,

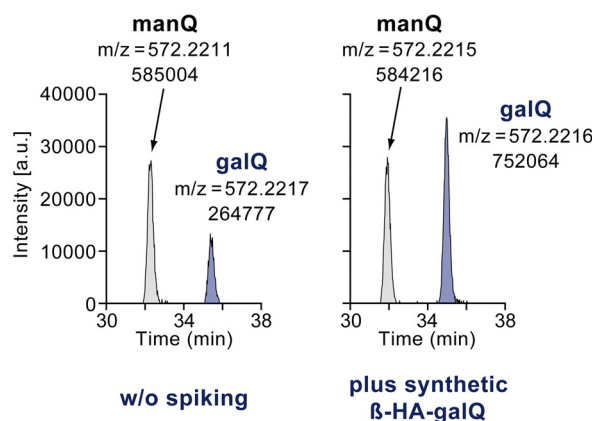


Figure 3. Results of a co-injection study confirming the identity of our synthetic β -homoallylic-galQ **2** and the natural product. Depicted are the extracted-ion chromatograms ($m/z = 572.2148$ – 572.2248) of two HPLC-MS analyses, either with (right) or without (left) prior spiking of the synthetic galQ **2**. The exact mass $[M+H]^+$ of galQ (and manQ) is 572.2198 u, showing a perfect match to the two MS peaks observed by us with less than 4 ppm deviation.

kidney, heart, and spleen, using the same RNA isolation and digestion protocol as for the co-injection experiment (Figure 4A).

From our data it is clearly evident that the levels of all three modifications (galQ, manQ, and Q) generally increase with age. This effect is by far most pronounced with Q, while galQ and manQ only show a modest increase, if at all. Furthermore, and for all three modifications investigated, we see differences between the six organs at the same age. Heart, followed by brain tissues, contains the largest levels of Q and its sugar-modified derivatives, followed by kidney, liver, and spleen. In general, the changes of the modification levels observed by us positively correlate with the respective organ-specific protein-synthesis demands, as we have shown before.^[7] Nevertheless, there are some prominent outliers. These outliers (e.g. heart tissue) seem to rather correlate with the organ-specific density of mitochondria. It was shown before that the Q-base in mitochondrial tRNA^{Tyr} and tRNA^{Asp} is not sugar-modified.^[8] We therefore speculate that the organ-specific differences in the levels of galQ, manQ, and Q are due to a combination of two independent effects: The organ-specific protein-synthesis ratio and the organ-specific mitochondrial density.

It is well-established that for biosynthesis of Q (and its sugar-modified derivatives), eukaryotes have to take up the queuine base from their diet,^[21,22] mammals thereby profiting from their gut microbiome.^[23] We therefore speculate that the low levels of Q-family nucleosides in newborn mice observed here may be caused by a lack of queuine supply in newborn mice, which only later establish their microbiome. Furthermore, high rates of cell division and tissue development in young mice may cause additional queuine supply problems.

To further study the influence of queuine availability on Q-family modification levels, cell culture experiments were performed: Human embryonic kidney cells (HEK 293T) were grown either in culture medium supplemented with 20 nM queuine (enriched medium) or in medium without additional

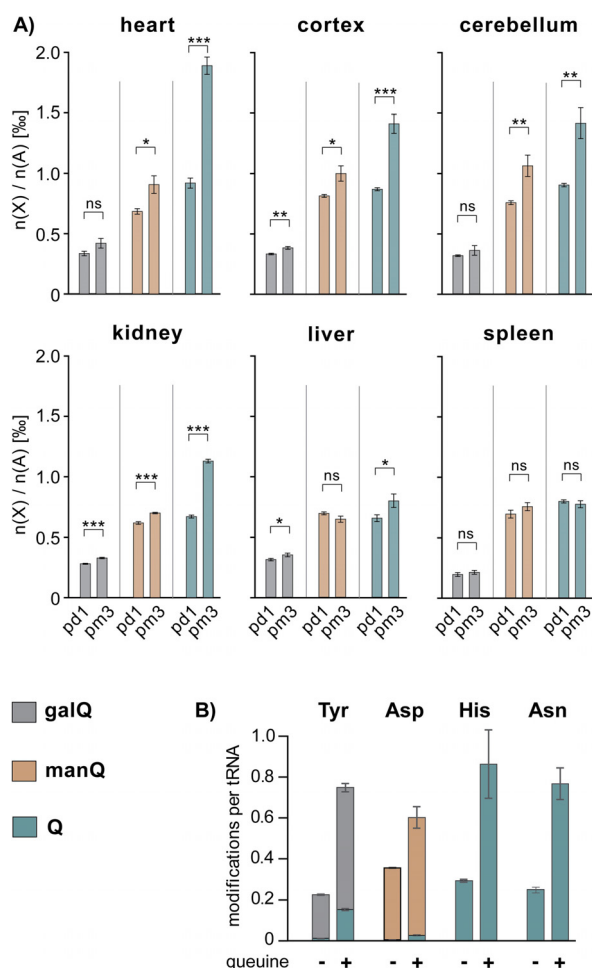


Figure 4. A) Absolute levels of galQ **2**, manQ **3**, and Q **1** in six different organs of newborn (postnatal day 1, pd1) and adult (postnatal month 3, pm3) mice. Values are given as number of xQ modifications $n(X)$ per 1000 adenosine nucleosides $n(A)$. Error bars represent the standard deviation of three biological replicates. For statistical analysis, Student's unpaired two-tailed t-test was used. n.s.: not significant, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$. B) Number of galQ, manQ, and Q modifications per cytosolic $tRNA^{Tyr}$, $tRNA^{Asp}$, $tRNA^{His}$, and $tRNA^{Asn}$ from human HEK 293T cells, respectively. queueine+: cells grown in queueine-enriched medium, queueine-: cells grown in standard medium. Values are given as average number of modifications per tRNA molecule. Error bars represent the standard deviation of three biological replicates.

queueine (standard medium). Queueine is the substrate of the TGT enzyme, which performs the exchange of a guanine base by the queueine heterocycle during tRNA maturation.^[24–26] From both cell populations, cytosolic $tRNA^{Tyr}$, $tRNA^{Asp}$, $tRNA^{His}$, and $tRNA^{Asn}$ were isolated and digested to the nucleoside level. For each of these four individual tRNA species, the number of galQ, manQ, and Q modifications per tRNA was then determined by a mass-spectrometry-based isotope-dilution method using the reference compound synthesized here (see the Supporting Information).

Indeed, our data show that the extent of Q-modification in the wobble position of cytosolic $tRNA^{Tyr}$, $tRNA^{Asp}$, $tRNA^{His}$, and $tRNA^{Asn}$ is strongly dependent on queueine availability (Figure 4B).^[24] In the case of $tRNA^{Tyr}$ (galQ), $tRNA^{His}$, and

$tRNA^{Asn}$ (Q), the difference in modification extent between cells grown in enriched versus standard medium is threefold, while for $tRNA^{Asp}$ (manQ) it is 1.7 fold. These results are well in line with our hypothesis and might therefore explain the lower modification levels in newborn mice. Of note, in all of our experiments even a sufficient queueine supply did not lead to fully modified tRNAs. This might again be an indication of the modification machinery lagging behind the de novo synthesis of tRNA in highly proliferating cells.

Furthermore, we detected a Q-only-modified $tRNA^{Tyr}$ form in our experiments lacking the galactose sugar, while $tRNA^{Asp}$ was always found to be either modified with manQ or completely unmodified. It seems that, in our experimental setup, mannosylation of $tRNA^{Asp}$ may be more tightly connected to G/Q-exchange than the galactosylation of Q-only-bearing $tRNA^{Tyr}$. Testing this exciting hypothesis is an interesting starting point for future studies.

In summary, we here report the first total synthesis of the human natural product galactosyl-queuosine **2**. Our synthetic material allowed us to confirm the proposed galQ structure by direct comparison with natural material, and we show that this hypermodified nucleoside is present in all tissues of newborn and adult mice. We furthermore report the absolute levels of all three Q-family members in six different mouse organs and in human cytosolic tRNAs. Taken together, our results confirm the crucial importance of tRNA galQ and manQ modification.

Acknowledgements

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) -Project-ID 255344185 – SPP 1784; Project-ID 325871075 – SFB 1309, and Project-ID 393547839 – SFB 1361. This project has received additional funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 765266 (LightDyNAMics) and from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement No 741912 (EPiR).

Conflict of interest

The authors declare no conflict of interest.

Keywords: galactosylation · mannosylation · mannosyl-queuosine · queuosine · RNA modifications

- [1] P. Boccaletto, M. A. Machnicka, E. Purta, P. Piątkowski, B. Bagiński, T. K. Wirecki, V. de Crécy-Lagard, R. Ross, P. A. Limbach, A. Kotter, M. Helm, J. M. Bujnicki, *Nucleic Acids Res.* **2018**, *46*, D303–D307.
- [2] F. Harada, S. Nishimura, *Biochemistry* **1972**, *11*, 301–308.
- [3] H. Kasai, Y. Kuchino, K. Nihei, S. Nishimura, *Nucleic Acids Res.* **1975**, *2*, 1931–1940.
- [4] H. Grosjean, M. Sprinzl, S. Steinberg, *Biochimie* **1995**, *77*, 139–141.

- [5] A. Costa, J.-P. Pais de Barros, G. Keith, W. Baranowski, J. Desgrès, *J. Chromatogr. B* **2004**, *801*, 237–247.
- [6] D. Globisch, D. Pearson, A. Hienzsch, T. Brückl, M. Wagner, I. Thoma, P. Thumbs, V. Reiter, A. C. Kneuttinger, M. Müller, S. A. Sieber, T. Carell, *Angew. Chem. Int. Ed.* **2011**, *50*, 9739–9742; *Angew. Chem.* **2011**, *123*, 9913–9916.
- [7] C. Brandmayr, M. Wagner, T. Brückl, D. Globisch, D. Pearson, A. C. Kneuttinger, V. Reiter, A. Hienzsch, S. Koch, I. Thoma, P. Thumbs, S. Michalakakis, M. Müller, M. Biel, T. Carell, *Angew. Chem. Int. Ed.* **2012**, *51*, 11162–11165; *Angew. Chem.* **2012**, *124*, 11324–11328.
- [8] T. Suzuki, T. Suzuki, *Nucleic Acids Res.* **2014**, *42*, 7346–7357.
- [9] N. Okada, N. Shindo-Okada, S. Nishimura, *Nucleic Acids Res.* **1977**, *4*, 415–423.
- [10] N. Okada, S. Nishimura, *Nucleic Acids Res.* **1977**, *4*, 2931–2937.
- [11] H. Kasai, K. Nakanishi, R. D. Macfarlane, D. F. Torgerson, Z. Ohashi, J. A. McCloskey, H. J. Gross, S. Nishimura, *J. Am. Chem. Soc.* **1976**, *98*, 5044–5046.
- [12] T. Ohgi, T. Kondo, T. Goto, *J. Am. Chem. Soc.* **1979**, *101*, 3629–3633.
- [13] T. Kondo, K. Okamoto, T. Ohgi, T. Goto, *Tetrahedron* **1986**, *42*, 207–213.
- [14] F. Klepper, E.-M. Jahn, V. Hickmann, T. Carell, *Angew. Chem. Int. Ed.* **2007**, *46*, 2325–2327; *Angew. Chem.* **2007**, *119*, 2377–2379.
- [15] A. M. Szpilman, E. M. Carreira, *Org. Lett.* **2009**, *11*, 1305–1307.
- [16] D. Crich, T. J. Ritchie, *Carbohydr. Res.* **1989**, *190*, C3–C6.
- [17] U. P. Singhi, R. K. Brown, *Can. J. Chem.* **1971**, *49*, 1179–1186.
- [18] D. Vonlanthen, C. J. Leumann, *Synthesis* **2003**, 1087–1090.
- [19] H. Ovaa, B. Lastdrager, J. D. C. Codée, G. A. van der Marel, H. S. Overkleeft, J. H. van Boom, *J. Chem. Soc. Perkin Trans. 1* **2002**, 2370–2377.
- [20] A. M. Szpilman, J. M. Manthorpe, E. M. Carreira, *Angew. Chem. Int. Ed.* **2008**, *47*, 4339–4342; *Angew. Chem.* **2008**, *120*, 4411–4414.
- [21] J. P. Reyniers, J. R. Pleasants, B. S. Wostmann, J. R. Katze, W. R. Farkas, *J. Biol. Chem.* **1981**, *256*, 11591–11594.
- [22] J. R. Katze, B. Basile, J. A. McCloskey, *Science* **1982**, *216*, 55–56.
- [23] W. R. Farkas, *J. Biol. Chem.* **1980**, *255*, 6832–6835.
- [24] N. Okada, F. Harada, S. Nishimura, *Nucleic Acids Res.* **1976**, *3*, 2593–2604.
- [25] N. Shindo-Okada, N. Okada, T. Ohgi, T. Goto, S. Nishimura, *Biochemistry* **1980**, *19*, 395–400.
- [26] C. Boland, P. Hayes, I. Santa-Maria, S. Nishimura, V. P. Kelly, *J. Biol. Chem.* **2009**, *284*, 18218–18227.
- [27] F. Tuorto, C. Legrand, C. Cirzi, G. Federico, R. Liebers, M. Müller, A. E. Ehrenhofer-Murray, G. Dittmar, H.-J. Gröne, F. Lyko, *EMBO J.* **2018**, *37*, e99777.

Manuscript received: February 14, 2020

Accepted manuscript online: March 11, 2020

Version of record online: April 21, 2020

6. Unveröffentlichte Arbeiten

6.1. Weitere 6-Aza-substituierte Analoga der epigenetisch relevanten dC-Modifikationen

6.1.1. Synthesen und Eigenschaften von 6-Aza-mdC und 6-Aza-hmdC

Zur Weiterführung der mechanistischen Studien von 6-Aza-fdC und dessen potentiellen Deformylierungsproduktes 6-Aza-dC (Kapitel 5.1), sollten auch die Effekte der 6-Aza-substituierten Nukleosid-Analoga 6-Aza-mdC und 6-Aza-hmdC bei Verfütterungsexperimenten an Zellen untersucht werden (siehe Überblick in Abbildung 6-1).

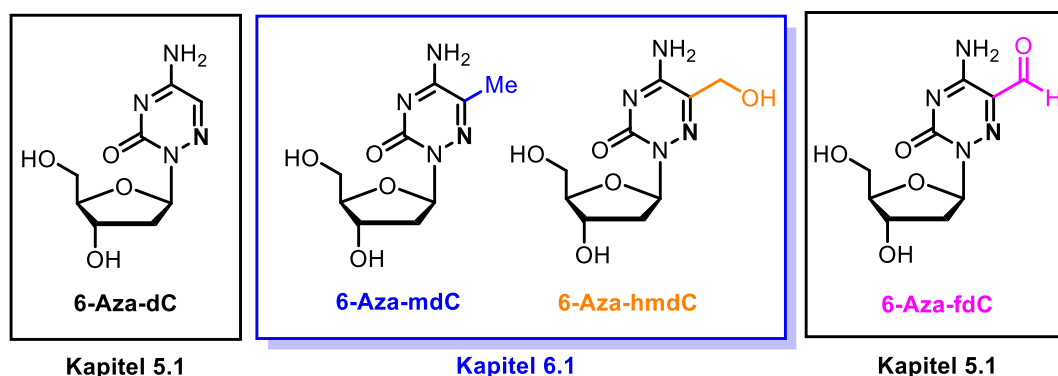


Abbildung 6-1: Überblick über die untersuchten 6-Aza-2'-desoxycytidinderivate und zugehörige Kapitel-einordnung.

Die erste untersuchte Route zur Synthese von 6-Aza-mdC erfolgte in Analogie zu den etablierten Synthesen von F-mdC über eine *Kumada*-Kreuzkupplungsreaktion des an 5-Position iodierten Nukleosids **3** mit Trimethylaluminium (Abbildung 6-2).^[61] Interessanterweise war dabei zu beobachten, dass bei der Iodierung der 5-Position von 6-Aza-2'-desoxyuridin **2** unter den bekannten Bedingungen zur oxidativen Iodierung von Pyrimidinen keinerlei Produktbildung auftrat. Der zusätzliche Ringstickstoff an der 6-Position schien dabei auf Grund seines elektronenziehenden Effekts die Reaktivität der 5-Position zu verringern. Eine Iodierung des Nukleosids konnte auch nicht bei verlängerter Reaktionsdauer oder erhöhter Temperatur erreicht werden. Allerdings konnte in Anlehnung an bekannte Verfahren eine Iodierung des freien Heterozyklus 6-Azauracil **1** unter sehr harschen Reaktionsbedingungen und mit einer sehr schlechten Ausbeute durchgeführt werden.^[105,106] Auch die anschließende *Vorbrüggen*-Nukleosidierung lieferte keine guten Ausbeuten, nichtsdestotrotz konnte ein anomeres Gemisch von α - und β -5-Iodo-6-aza-2'-desoxyuridin **3** erhalten werden, welches durch Umkristallisation in siedendem

Ethanol aufgetrennt werden konnte. Eine anschließende palladiumkatalysierte *Kumada*-Kreuzkupplung mit AlMe_3 unter den für Pyrimidine optimierten Bedingungen^[61] lieferte mit dem 6-Aza-Nukleosid **3** allerdings ebenfalls kein gewünschtes Produkt. Da der zusätzliche Ringstickstoff an der 6-Position der 6-Aza-Nukleoside einen so deutlichen Effekt auf die Reaktivität der 5-Position zu haben schien, wurde im weiteren Verlauf auf Substitutionsreaktionen an der 5-Position verzichtet und eine andere Syntheseroute gewählt.

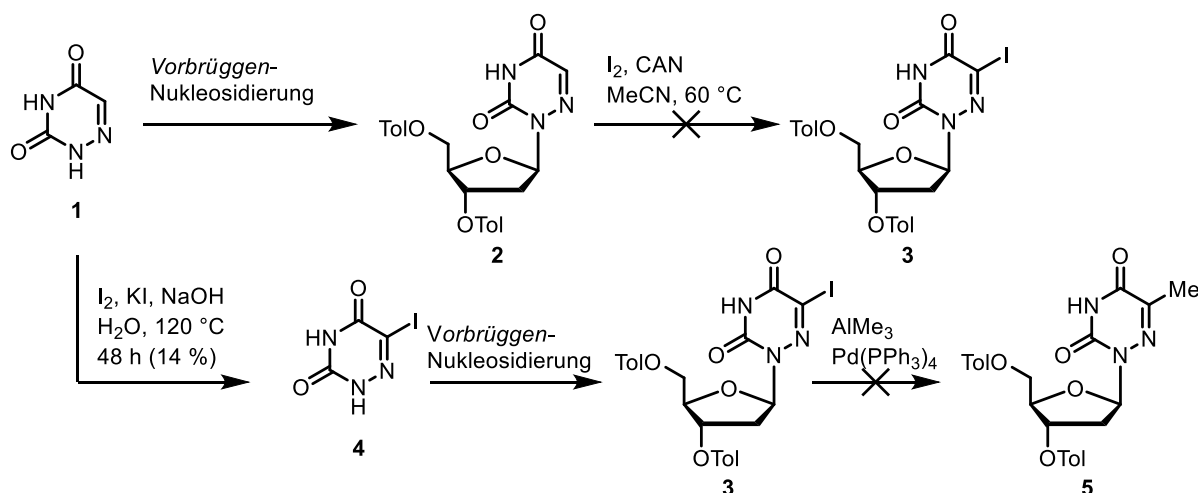


Abbildung 6-2: Untersuchte Syntheserouten über Iodierung des Heterozyklus.

Eine Synthese des bereits methylierten Heterozyklus vor der Nukleosidierung ist bekannt und kann unter anderem in Analogie zu der Synthese von 6-Aza-5-Hydroxymethyluracil (siehe Kapitel 5.1) über die Zyklisierung eines geeigneten Semicarbazons durchgeführt werden.^[107] Allerdings ist die gewünschte Verbindung auch kommerziell erhältlich. Ausgehend von 6-Azathymin **6** konnte eine *Vorbrüggen*-Nukleosidierung mit *Hoffer's* Chlorzucker in guter Ausbeute durchgeführt werden. Das gewünschte β -Nukleosid war in einem deutlichen Überschuss im resultierenden anomeren Gemisch enthalten. Durch Umkristallisation des Anomerengemisches in siedendem Ethanol konnte das isomerenreine β -Nukleosid erhalten werden. Auch die anschließende Aminierung der 4-Position über eine Substitution des Triazolid-Intermediats lieferte, wie bei der Aminierung der anderen 6-Aza-Nukleosidderivate, eine hohe Ausbeute. Die 4-Position scheint hierbei durch die elektronischen Effekte des Ringstickstoffs der 6-Position für derartige Reaktionen grundsätzlich aktiviert zu sein. Entschützung von den Toluoyl-Schutzgruppen an 3'- und 5'-Positionen mit ethanolischem Methylamin und anschließende chromatographische Aufreinigung mittels semipräparativer HPLC lieferte das gewünschte Zielmolekül **9** in guter

Ausbeute und ausreichender Reinheit für zellbiologische Experimente (siehe Abbildung 6-3).

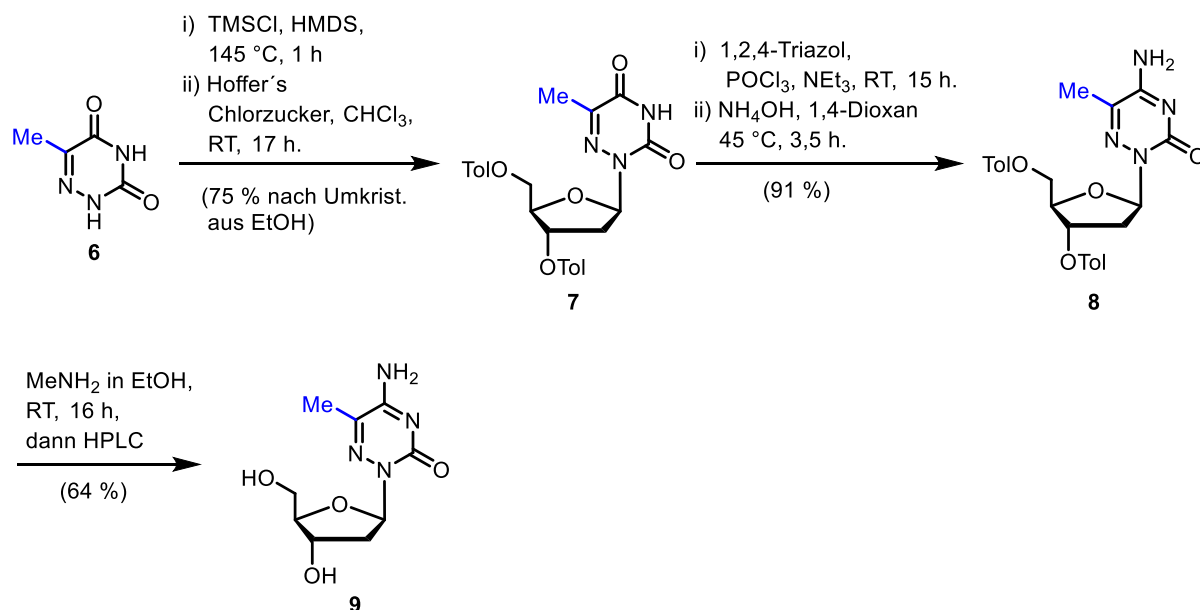


Abbildung 6-3: Syntheseroute ausgehend von 6-Azathymine 6.

Zur Synthese von 6-Aza-hmdC **13** sind bereits Vorschriften bekannt,^[108,109] allerdings umfassen viele der Stufen veraltete, sehr langwierige Methoden oder stark toxisches Material. Die Nukleosidierung beispielsweise wird hierbei unter *Koenigs-Knorr*-Bedingungen mit Quecksilberacetat in Benzol durchgeführt und das Produkt schließlich über einen längeren Zeitraum in einer mäßigen Ausbeute von 41% auskristallisiert. Ein weiteres Beispiel in der Literatur ist die Aminierung eines geschützten 6-Aza-2'-desoxyuridinderivates mit anschließender Entschützung zum 6-Aza-hmdC, die zur Isolation des Produkts eine Wartezeit von zwei Wochen beschreibt.^[108] Um solche Reaktionsbedingungen und -zeiten zu vermeiden, wurde die in dieser Arbeit gezeigte Synthese von 6-Aza-hmdC **13** auf Basis modernerer Verfahren entwickelt. Diese basiert auf der neu entwickelten Synthese von 6-Aza-fdC (Kapitel 5.1 und Abbildung 6-4). Zur Darstellung von 6-Aza-hmdC **13** wird ein 4-Triazolid-Intermediat zur Aminierung von **12** direkt in einem Eintopfverfahren in methanolischem Ammoniak an der 4-Position aminierte und gleichzeitig entschützt. Dadurch konnte direkt das Zielmolekül 6-Aza-hmdC **13** erhalten werden, das nach Aufreinigung über semipräparative HPLC für die Entwicklung einer Methode zur Detektion der Verbindung über *Triple-quadrupol*-(TQ)-Massenspektrometrie verwendet werden konnte.

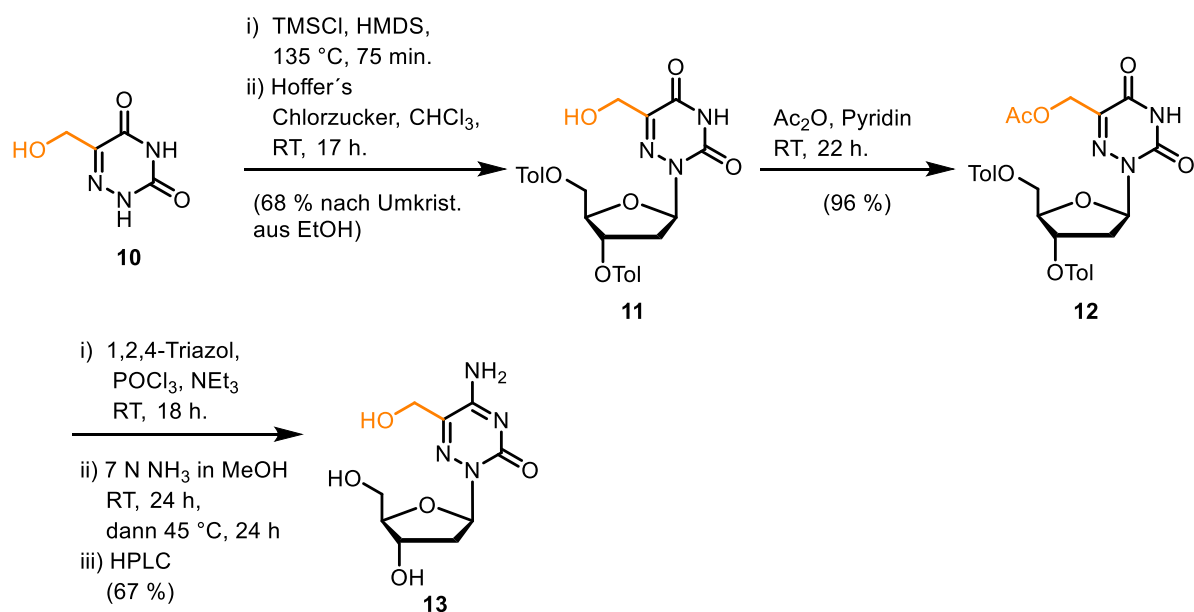


Abbildung 6-4: Schema der Synthese von 6-Aza-hmdC **13**.

Ein UHPL-Chromatogramm, das zusammenfassend die verschiedenen synthetisierten 6-Aza-Nukleoside bei ihrem jeweiligen Übergang bei der TQ-Massenspektrometrie zeigt, ist in Abbildung 6-5 dargestellt.

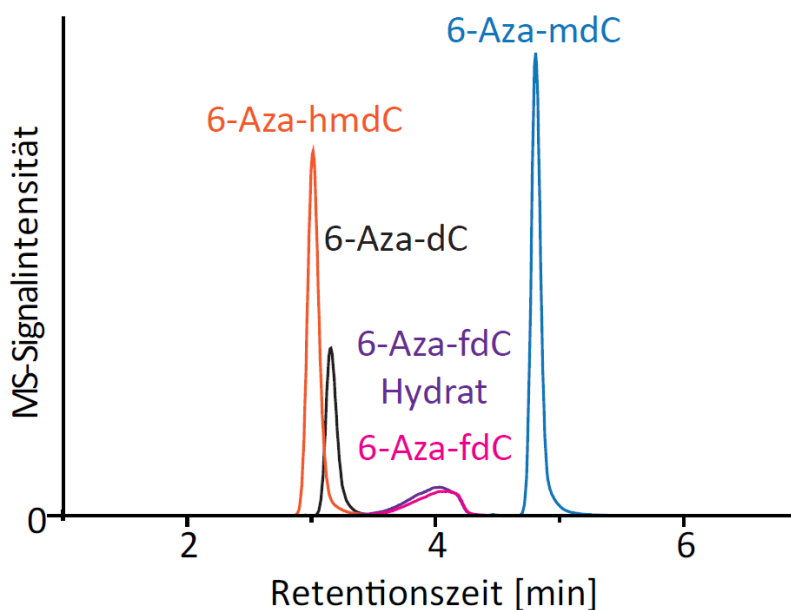


Abbildung 6-5: UHPL-Chromatogramm der verschiedenen, synthetisierten 6-Aza-dC-Derivate.

Anhand des UHPL-Chromatogramms ist ersichtlich, dass 6-Aza-mdC und 6-Aza-hmdC vergleichbar mit 6-Aza-dC und wie erwartet ein scharfes Elutionsprofil aufweisen, wohingegen 6-Aza-fdC auf Grund seiner Tendenz zur Hydratbildung ein breites Elutionsprofil zeigt, weswegen eine Derivatisierung dieser Verbindung zur effizienten Detektion notwendig wurde (Kapitel 5.1). Die Entwicklung der Methoden

zur massenspektrometrischen Detektion der Verbindungen mittels hochsensitiver TQ-Massenpektrometrie wurde von *Dr. Sarah Schiffers* durchgeführt und Details sind ihrer Dissertation zu entnehmen.^[110] Im Nachfolgenden soll nur ein sehr kurzer Überblick über die Ergebnisse der massenpektrometrischen Untersuchungen der Zellverfütterungsexperimente von 6-Aza-mdC, die von *Ewelina Kaminska* durchgeführt wurden, gegeben werden.

Die bei der TQ-massenspektrometrischen Analyse detektierten Übergänge von 6-Aza-mdC und 6-Aza-hmdC sind in Abbildung 6-6 dargestellt.

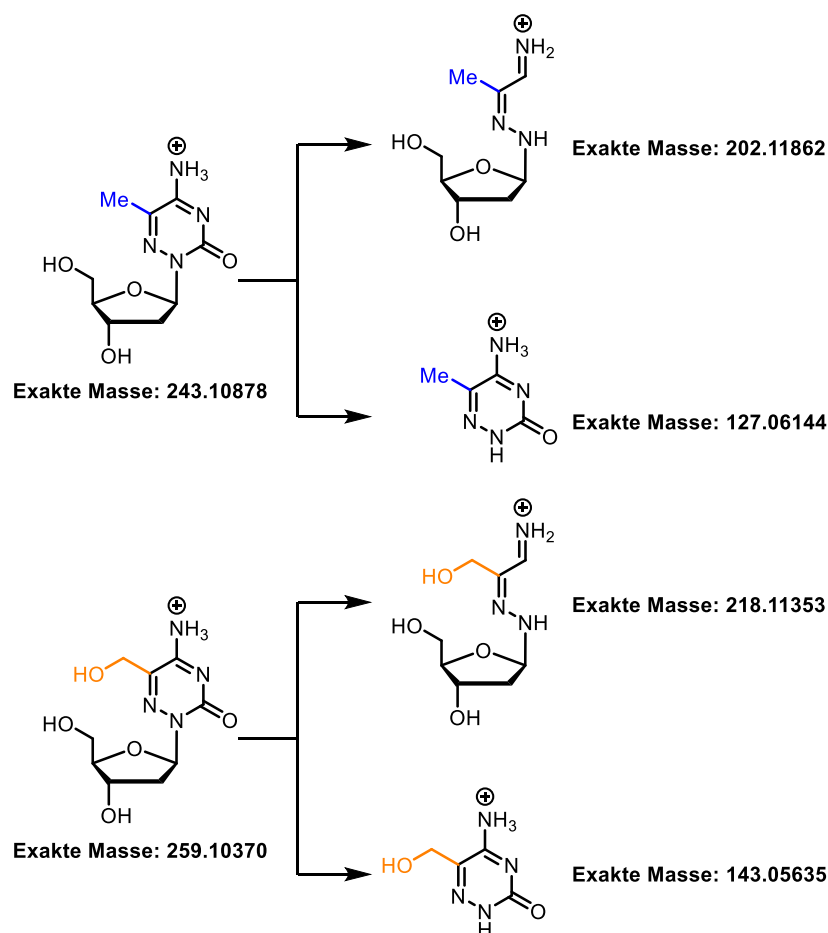


Abbildung 6-6: Schema zur ESI-massenspektrometrischen Fragmentierung von 6-Aza-mdC und 6-Aza-hmdC; exakte Massen wurden zur Veranschaulichung berechnet.

Beim Vergleich des Fragmentationsverhaltens der beiden Verbindungen im MS/MS ist ersichtlich, dass beide Verbindungen sehr ähnliche Fragmentierungsmuster zeigen. Die prominentesten Massenübergänge durch Fragmentierung sind bei beiden Verbindungen entweder eine Abspaltung einer Formamid-Einheit aus dem Heterozyklus oder die Spaltung der glykosidischen Bindung, wobei hier die Spaltung

der glykosidischen Bindung den am besten zu detektierenden Massenübergang darstellt und daher zur Detektion dieser Moleküle herangezogen wurde.

Auf Basis der synthetisierten Verbindungen 6-Aza-mdC und 6-Aza-hmdC und der erfolgreichen Entwicklung der Methoden zu ihrer Detektion sollten nun Zellverfütterungsexperimente mit 6-Aza-mdC durchgeführt werden. Das Ziel dieser Studien war zu ermitteln, ob nach erfolgreichem Einbau des verfütterten 6-Aza-mdCs dieses im Genom durch die TET-Enzyme zum 6-Aza-hmdC oder sogar weiter bis hin zum 6-Aza-fdC oxidiert werden kann. Da bekannt ist, dass der Mechanismus der TET-Enzyme nicht über einen nukleophilen Angriff an die 6-Position der Nukleobase verläuft,^[111–113] sollte die 6-Aza-Substitution keinen direkten Effekt auf die Aktivität der TET-Enzyme an 6-Aza-mdC haben. Ein Nachweis der TET-katalysierten Oxidation der 6-Aza-Nukleosidderivate könnte im weiterführenden Rahmen der in Kapitel 5.1 dargelegten Studien verwendet werden, um genomisches 6-Aza-fdC durch Zellverfütterung von 6-Aza-mdC und nicht durch direkte Fütterung von 6-Aza-fdC einzuführen.

Bei Zellverfütterungsexperimenten von 6-Aza-mdC an mESCs unter dem Toxizitätslimit der Verbindung konnte in der anschließenden massenspektrometrischen Analyse der verdauten DNA-Extrakte kein Einbau von 6-Aza-mdC detektiert werden. Darüber hinaus wurden entsprechend auch dessen potentielle Oxidationsprodukte 6-Aza-hmdC und 6-Aza-fdC bei diesen Untersuchungen nicht detektiert. Hochaufgelöste massenspektrometrische Untersuchungen, die von *Dr. Mirko Wagner* durchgeführt wurden, zeigten ein Signal, das 6-Aza-2'-desoxythymidin (6-Aza-dT), dem potentiellen Desaminierungsprodukts von 6-Aza-mdC, zugeordnet werden kann. Dieses Ergebnis konnte allerdings nicht verifiziert werden und in allen weiteren Studien konnten darüber hinaus keine Desaminierungsprodukte der anderen 6-Aza-Nukleoside nachgewiesen werden. Da das detektierte Signal äußerst geringe Intensität aufwies, wurden keine weiteren Experimente zur Zellverfütterung von 6-Aza-mdC durchgeführt.

Zusammenfassend kann festgestellt werden, dass zwar die Synthese von 6-Aza-mdC und 6-Aza-hmdC auf Grundlage einer neu entwickelten Syntheseroute und die Detektion der Verbindungen über hochsensitive TQ-Massenspektrometrie gelang. Allerdings konnten bei den anschließend durchgeführten Verfütterungsexperimenten

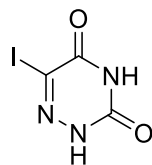
auf Grund der nicht erfolgten Inkorporation von 6-Aza-mdC in das Genom der Zellen keine weiteren Informationen erhalten werden.

6.1.2. Projektbeiträge

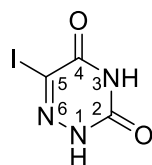
Im Rahmen dieses Projekts entwickelte ich die Synthesen der 6-Aza-substituierten Nukleoside 6-Aza-mdC und 6-Aza-hmdC und setzte diese um. *Ewelina Kaminska* führte die Zellverfütterungsexperimente durch und *Dr. Sarah Schiffers* entwickelte die Methoden zur Detektion der Verbindungen mittels TQ-Massenspektrometrie und führte DNA-Verdaus sowie TQ-massenspektrometrische Messungen der Proben aus der Zellkultur durch. *Dr. Mirko Wagner* führte hochaufgelöste massenspektrometrische *Orbitrap*-Untersuchungen durch.

6.1.3. Experimentaltteil

5-Iodo-6-azauracil (**4**)



Die Synthese erfolgte gemäß einer modifizierten Vorschrift von *Kögler et al.*^[106] In einem Rundkolben wurden 2.65 g 6-Azauracil **1** (23.4 mmol, 1.0 eq.) in 82 mL ddH₂O suspendiert und mit elementarem Iod (11.70 g, 46.1 mmol, 2.0 eq.), Kaliumiodid (12.39 g, 74.6 mmol, 3.2 eq.) und frisch gepulvertem NaOH (3.74 g, 93.6 mmol, 4.0 eq.) versetzt. Das Gemisch wurde 48 h refluxiert (Ölbadtemperatur 125 °C), anschließend auf 0 °C gekühlt und mit konz. HCl auf einen pH von 1 eingestellt. Daraufhin wurden 60 mL einer 5%igen wässrigen NaHSO₃/Na₂S₂O₅-Lösung unter Ausbildung einer gelben Lösung zugegeben. Diese wurde mit EtOAc extrahiert (3 x 300 mL), die vereinigten organischen Phasen wurden mit einem Gemisch aus 20 mL einer 5%igen wässrigen NaHSO₃/Na₂S₂O₅-Lösung und 110 mL ddH₂O gewaschen, über Na₂SO₄ getrocknet und *in vacuo* zur Trockene eingengt. Der Rückstand wurde mittels Säulenchromatographie an Silicagel (2% EtOH in CH₂Cl₂) aufgereinigt, um 794 mg der iodierten Verbindung **4** (3.3 mmol, 14%) als gelblichen Feststoff zu liefern.



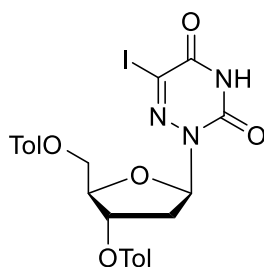
¹H-NMR (600 MHz, DMSO-*d*₆): δ/ppm = 12.60 (s, 1H, NH), 12.14 (s, 1H, NH).

¹³C-NMR (151 MHz, DMSO-*d*₆): δ/ppm = 155.05 (4-C), 148.85 (2-C), 110.62 (5-C).

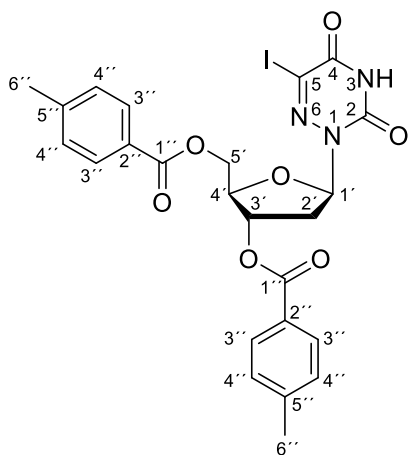
HRMS (ESI⁻): *m/z* berechnet für C₃HN₃O₂I⁻ [M-H]⁻ 237.9119; gefunden: 237.9120.

R_f (CH₂Cl₂:EtOH 9:1 mit 1% HCOOH): 0.45.

5-Iodo-(3',5'-di-O-p-toluoyl)-6-aza-2'-desoxyuridin (**3**)



In einem ausgeheizten Rundkolben wurden 205 mg 5-Iodo-6-azauracil **4** (1.05 mmol, 1.0 eq.) mit 1.14 mL Hexamethyldisilazan (HMDS) versetzt und der Suspension 89 μ L Trimethylchlorsilan (TMSCl) zugegeben. Anschließend wurde das Gemisch unter Argonatmosphäre für 4 h auf 135 °C erwärmt, anschließend auf Raumtemperatur kommen lassen und mit trockenem CHCl_3 (3 x 5 mL) coevaporiert. Der ölige Rückstand wurde am Hochvakuum getrocknet, daraufhin in 3.30 mL trockenem CHCl_3 gelöst, mit 449 mg Hoffer's Chlorzucker (1.16 mmol, 1.1 eq.) versetzt und 16 h bei Raumtemperatur verrührt. Das Reaktionsgemisch wurde dann auf gesättigte wässrige NaHCO_3 -Lösung (50 mL) gegossen und mit CH_2Cl_2 (2 x 50 mL) extrahiert. Die vereinigten organischen Phasen wurden mit gesättigter wässriger Natriumchloridlösung (50 mL) gewaschen, über Na_2SO_4 getrocknet und *in vacuo* zur Trockene eingengt. Säulenchromatographische Aufreinigung des Rohprodukts an Silicagel (iHex:EtOAc 3:1 \rightarrow iHex:EtOAc 2:1 \rightarrow iHex:EtOAc 1:1) lieferte ein Gemisch aus α - und β -Nukleosidierungsprodukten (ca. 40% α -Nukleosid via NMR), das durch Umkristallisation in siedendem EtOH aufgetrennt werden konnte, um 223 mg des β -Nukleosids **3** (0.38 mmol, 36%) in Form eines weißen Feststoffs zu liefern.

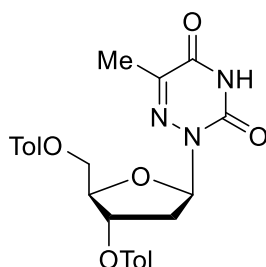


¹H-NMR (600 MHz, CDCl₃): δ/ppm = 8.70 (s, 1H, NH), 7.98 – 7.87 (m, 4H, 3''-H), 7.28 – 7.15 (m, 4H, 4''-H), 6.61 (t, *J* = 6.4 Hz, 1H, 1'-H), 5.67 (dt, *J* = 6.3, 3.0 Hz, 1H, 3'-H), 4.58 – 4.45 (m, 3H, 4'-H, 5'-H), 3.00 (dt, *J* = 14.2, 6.4 Hz, 1H, 2'-H), 2.53 (ddd, *J* = 14.2, 6.7, 3.7 Hz, 1H, 2'-H), 2.43 (s, 3H, 6''-H), 2.40 (s, 3H, 6''-H).

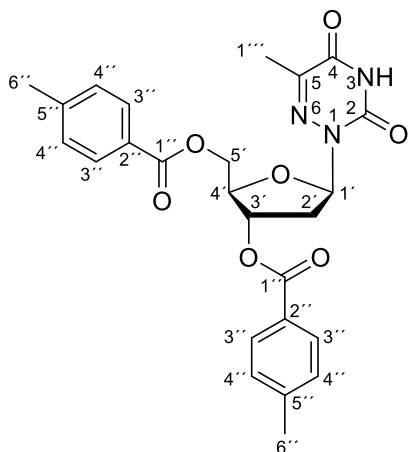
R_f (CH₂Cl₂:MeOH 20:1): 0.30.

Weitere analytische Daten sind bei *Kögler et al.* zu finden.^[106]

5-Methyl-(3',5'-di-O-p-toluoyl)-6-aza-2'-desoxyuridin (**7**)



Unter Argonatmosphäre wurden 300 mg 6-Azathymin **6** (2.36 mmol, 1.0 eq.) in 2.56 mL HMDS suspendiert und mit 89 µL TMSCl versetzt. Das Gemisch wurde 1 h bei 145 °C unter Rückfluss erhitzt und anschließend wurde die resultierende Lösung nach Abkühlen auf Raumtemperatur *in vacuo* bei 30 °C eingengt. Das erhaltene klare Öl wurde 1 h unter Hochvakuum getrocknet und daraufhin wurde der Rückstand in 7.41 mL trockenem CHCl₃ gelöst. 1.01 g Hoffer's Chlorzucker (2.60 mmol, 1.1 eq.) wurden zugegeben und das Reaktionsgemisch wurde 17 h bei Raumtemperatur verrührt. Daraufhin wurde das Reaktionsgemisch in 250 mL gesättigte wässrige NaHCO₃-Lösung gegossen und mit EtOAc (3 x 200 mL) extrahiert. Die vereinigten organischen Phasen wurden über Na₂SO₄ getrocknet und das Lösungsmittel *in vacuo* entfernt. Eine Aufreinigung mittels Säulenchromatographie an Silicagel (iHex:EtOAc 3:1 → iHex:EtOAc 2:1 → iHex:EtOAc 1:1) lieferte ein Gemisch der diastereomeren α- und β-Nukleosidierungsprodukte. Umkristallisation aus siedendem EtOH lieferte 852 mg des reinen β-Nukleosids **7** (1.78 mmol, 75%) als farblosen Feststoff.



$^1\text{H-NMR}$ (600 MHz, CDCl_3): δ/ppm = 8.75 (s, 1H, 3-H), 8.01 – 7.87 (m, 4H, 3''-H), 7.27 – 7.20 (m, 4H, 4'-H), 6.68 (t, J = 6.5 Hz, 1H, 1'-H), 5.69 (ddd, J = 6.6, 3.6, 3.3 Hz, 1H, 3'-H), 4.63 – 4.44 (m, 3H, 4'-H, 5'-H), 3.00 (dt, J = 13.9, 6.5 Hz, 1H, 2'-H), 2.46 (ddd, J = 14.1, 6.7, 3.7 Hz, 1H, 2'-H), 2.43 (s, 3H, 6''-H), 2.40 (s, 3H, 6''-H), 2.18 (s, 3H, 1'''-H).

$^{13}\text{C-NMR}$ (151 MHz, CDCl_3): δ/ppm = 166.38 (1''-C), 166.16 (1''-C), 155.80 (4-C), 148.51 (2-C), 145.01 (5-C), 144.51 (5''-C), 144.08 (5''-C), 129.92 (3''-C), 129.91 (3''-C), 129.37 (4''-C), 129.26 (4''-C), 127.03 (2''-C), 126.61 (2''-C), 85.78 (1'-C), 82.35 (4'-C), 75.12 (3'-C), 64.35 (5'-C), 35.05 (2'-C), 21.90 (6''-C), 21.86 (6''-C), 16.54 (1'''-C).

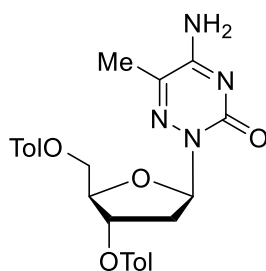
HRMS (ESI⁺): m/z berechnet für $\text{C}_{25}\text{H}_{26}\text{N}_3\text{O}_7^+$ $[\text{M}+\text{H}]^+$ 480.1765, gefunden: 480.1766; m/z berechnet für $\text{C}_{25}\text{H}_{29}\text{N}_4\text{O}_7^+$ $[\text{M}+\text{NH}_4]^+$ 497.2031, gefunden: 497.2029; m/z berechnet für $\text{C}_{25}\text{H}_{25}\text{N}_3\text{O}_7\text{Na}^+$ $[\text{M}+\text{Na}]^+$ 502.1585, gefunden: 502.1584.

HRMS (ESI⁻): m/z berechnet für $\text{C}_{25}\text{H}_{24}\text{N}_3\text{O}_7^-$ $[\text{M}-\text{H}]^-$ 478.1620; gefunden: 478.1628.

IR (ATR): ν (cm^{-1}) = 3192 (m), 3086 (w), 1714 (s), 1698 (s), 1608 (m), 1454 (w), 1407 (w), 1374 (w), 1295 (m), 1255 (s), 1176 (m), 1109 (s), 1083 (s), 1066 (s), 1048 (m), 1017 (m), 990 (w), 942 (m), 807 (m), 750 (s), 733 (m).

R_f (iHex:EtOAc 1:1): 0.57.

5-Methyl-(3',5'-di-O-p-toluoyl)-6-aza-2'-desoxycytidin (**8**)

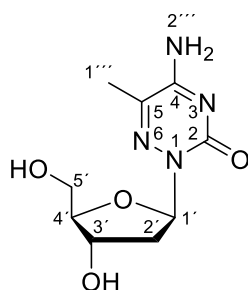


Unter Ar-Atmosphäre wurden 389 mg 1,2,4-Triazol (5.63 mmol, 9.0 eq.) in 15.80 mL trockenem MeCN gelöst und auf 0 °C gekühlt. 117 µL POCl₃ (1.25 mmol, 2.0 eq.) wurden tropfenweise zugegeben und das Gemisch wurde 10 min bei 0 °C verrührt. Anschließend wurden 772 µL NEt₃ (5.57 mmol, 8.9 eq.) zugetropft und das Gemisch für weitere 20 min bei 0 °C gerührt. Daraufhin wurden 300 mg des geschützten 5-Methyl-6-aza-2'-desoxyuridins **7** (1.07 mmol, 1.0 eq.) zugegeben. Das Reaktionsgemisch wurde auf Raumtemperatur erwärmen gelassen und 15 h bei dieser Temperatur verrührt. Nach vollständigem Umsatz des Startmaterials wurde das Gemisch in 150 mL einer gesättigten, wässrigen NaHCO₃-Lösung gegossen und mit CH₂Cl₂ (3 x 150 mL) extrahiert. Die vereinigten organischen Phasen wurden mit gesättigter wässriger Natriumchloridlösung (100 mL) gewaschen, über Na₂SO₄ getrocknet und Lösungsmittel wurden *in vacuo* entfernt.

Der Rückstand wurde in 12.60 mL 1,4-Dioxan gelöst, mit 3.97 mL konz. NH₄OH versetzt und 3.5 h bei 45 °C verrührt. Daraufhin wurde es in gesättigte wässrige NH₄Cl-Lösung (100 mL) gegossen und mit EtOAc (3 x 100 mL) extrahiert. Die vereinigten organischen Phasen wurden mit gesättigter wässriger Natriumchloridlösung (100 mL) gewaschen, über Na₂SO₄ getrocknet und *in vacuo* zur Trockene eingengt. Das Rohprodukt wurde mittels Säulenchromatographie an Silicagel (CH₂Cl₂:MeOH 40:1 → CH₂Cl₂:MeOH 20:1) aufgereinigt, um 275 mg des Aminierungsprodukts **8** (0.57 mmol, 91%) als gelben Feststoff zu erhalten.



8.0 mL einer ethanolischen Lösung von MeNH₂ (33 Gew.-%) wurden zu 120 mg der Toluoyl-geschützten Verbindung **8** (0.25 mmol, 1.0 eq.) gegeben und das Reaktionsgemisch wurde 16 h bei Raumtemperatur verrührt. Abschließend wurde das Gemisch *in vacuo* zur Trockene eingengt und der erhaltene Rückstand in 5 mL ddH₂O und 5 mL CH₂Cl₂ aufgenommen. Die Phasen wurden getrennt und die wässrige Phase weiter mit CH₂Cl₂ extrahiert (4 x 5 mL). Anschließend wurde die wässrige Phase *in vacuo* eingengt, das Rohprodukt in ddH₂O (18 mL) gelöst und mittels *reversed phase* HPLC (0 % → 5 % MeCN in H₂O in 45 min) aufgereinigt, um 39 mg des entschützten 6-Aza-Nukleosides **9** (0,16 mmol, 64%) als farblosen Feststoff zu liefern.



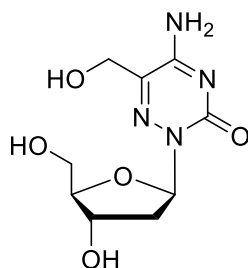
¹H-NMR (600 MHz, D₂O): δ/ppm = 6.46 (dd, *J* = 7.0, 5.0 Hz, 1H, 1'-H), 4.58 – 4.42 (m, 1H, 3'-H), 3.93 (dt, *J* = 6.1, 3.9, 1.7 Hz, 1H, 4'-H), 3.69 (ddt, *J* = 12.2, 3.8, 1.8 Hz, 1H, 5'-H), 3.59 (ddt, *J* = 12.3, 6.5, 1.8 Hz, 1H, 5'-H), 2.62 (ddd, *J* = 13.5, 6.7, 4.9 Hz, 1H, 2'-H), 2.25 – 2.19 (m, 4H, 2'-H, 1'''-H).

¹³C-NMR (151 MHz, D₂O): δ/ppm = 158.84 (4-C), 156.29 (2-C), 137.01 (5-C), 86.37 (4'-C), 86.17 (1'-C), 70.99 (3'-C), 61.80 (5'-C), 36.72 (2'-C), 16.25 (1'''-C).

HRMS (ESI⁺): *m/z* berechnet für C₉H₁₄N₄O₄Na⁺ [M+Na]⁺ 265.0907, gefunden: 265.0906.

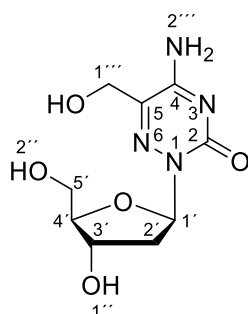
HRMS (ESI⁻): *m/z* berechnet für C₉H₁₃N₄O₄⁻ [M-H]⁻ 241.0942; gefunden: 241.0943.

5-Hydroxymethyl-6-aza-2'-desoxycytidin (**13**)



In einem ausgeheizten Schlenk-Kolben wurden 150 mg 1,2,4-Triazol (2.18 mmol, 9.0 eq.) unter Ar-Atmosphäre in 6.00 mL trockenem MeCN gelöst, auf 0 °C gekühlt, 45 µL POCl₃ (0.48 mmol, 2.0 eq.) wurden zugetropft und das Gemisch wurde 10 min bei 0 °C gerührt. Anschließend wurden 298 µL NEt₃ (2.15 mmol, 8.9 eq.) zugegeben und das Gemisch wurde weitere 20 min bei 0 °C verrührt. Daraufhin wurde das Reaktionsgemisch mit 130 mg des 2'-Desoxyuridin-Derivats **12** (0.24 mmol, 1.0 eq.) versetzt, auf Raumtemperatur erwärmen gelassen und 18 h bei dieser Temperatur verrührt. Die resultierende gelbe Suspension wurde in 100 mL gesättigte wässrige NaHCO₃-Lösung gegossen und mit CH₂Cl₂ (3 x 100 mL) extrahiert. Die vereinigten organischen Phasen wurden mit gesättigter wässriger Natriumchloridlösung (100 mL), gewaschen, über Na₂SO₄ getrocknet und Lösungsmittel wurden *in vacuo* entfernt.

Der erhaltene Rückstand wurde in 10 mL trockenem CH₂Cl₂ gelöst und auf zwei Aliquots aufgeteilt. Ein Aliquot wurde *in vacuo* zur Trockene eingengt und am Hochvakuum getrocknet. Der Rückstand wurde in 5.0 mL methanolischem NH₃ (7 N) gelöst und 24 h bei Raumtemperatur und weitere 24 h bei 45 °C verrührt. Nach vollständigem Umsatz des Startmaterials wurde das Reaktionsgemisch zur Trockene eingengt und am Hochvakuum getrocknet. Der gelbe Feststoff wurde in 10 mL ddH₂O und 10 mL CH₂Cl₂ gelöst, die Phasen wurden getrennt, die wässrige Phase wurde weiter mit CH₂Cl₂ (4 x 10 mL) extrahiert und schließlich lyophilisiert. Der Rückstand wurde in ddH₂O (12 mL) gelöst und mittels *reversed phase* HPLC (0 % → 5 % MeCN in H₂O in 45 min) aufgereinigt, um 20.5 mg des Zielmoleküls **13** (80 µmol, 67%) als farblosen Feststoff zu liefern.



¹H-NMR (600 MHz, D₂O): δ/ppm = 6.59 (dd, *J* = 7.0, 5.2 Hz, 1H, 1'-H), 4.67 (s, 2H, 1'''-H), 4.60 (dd, *J* = 6.0, 5.3 Hz, 1H, 3'-H), 4.05 (ddd, *J* = 6.5, 4.3, 2.1 Hz, 1H, 4'-H),

3.79 (dd, $J = 12.1, 4.1$, 1H, 5'-H), 3.69 (dd, $J = 12.1, 6.6$, 1H, 5'-H), 2.77 – 2.70 (m, 1H, 2'-H), 2.36 (dddd, $J = 13.2, 7.1, 5.4, 1.8$ Hz, 1H, 2'-H).

$^{13}\text{C-NMR}$ (151 MHz, D_2O): $\delta/\text{ppm} = 158.06$ (4-C), 156.05 (2-C), 136.82 (5-C), 86.59 (1'-C), 86.56 (4'-C), 71.07 (3'-C), 61.91 (5'-C), 59.99 (1'''-C), 36.82 (2'-C).

HRMS (ESI⁺): m/z berechnet für $\text{C}_9\text{H}_{15}\text{N}_4\text{O}_5^+$ $[\text{M}+\text{H}]^+$ 259.1037, gefunden: 259.1037; m/z berechnet für $\text{C}_9\text{H}_{14}\text{N}_4\text{O}_5\text{Na}^+$ $[\text{M}+\text{Na}]^+$ 281.0856, gefunden: 281.0856.

HRMS (ESI⁻): m/z berechnet für $\text{C}_9\text{H}_{11}\text{N}_4\text{O}_5^-$ $[\text{M}-\text{H}]^-$ 257.0891; gefunden: 257.0893.

R_f (CH_2Cl_2 :MeOH 5:1): 0.09.

6.2. Untersuchung der direkten Demethylierung mittels eines isotoopenmarkierten dC-Derivats

6.2.1. Synthese der Verbindung und erste Ergebnisse

Um eine aktive, direkte Demethylierung *in vivo* zu untersuchen, stellen Verfütterungsexperimente mit Verbindungen, die ein definiertes Isotoopenmarkierungsmuster aufweisen, einen zentralen Ansatz für interessante Experimente dar. Neben der Verfütterung eines isotoopenmarkierten fdC-Derivats^[59] bietet sich auch die Verfütterung isotoopenmarkierter dC-Derivate an. Ein solches kann dann nach Einbau in die DNA durch enzymatische Methylierung und Oxidation *in vivo* zu fdC umgewandelt und von dort auch wieder potentiell deformyliert werden. Dieser Ansatz bietet den Vorteil, dass markiertes fdC *in vivo* nur an den dafür im Genom vorgesehenen Positionen gebildet wird und in der Konsequenz in seinem natürlichen Verhalten untersucht werden kann. Bei einer Verfütterung eines fdC-Derivats kann dieses auch an dC-Positionen, die im biologischen Kontext kein fdC bilden würden, an Stelle eines dC eingebaut werden. Dieses könnte zu einer veränderten Prozessierung der fdCs an den nicht-biologisch relevanten dC-Positionen führen.

Dieses Projekt geht auf Studien von *Dr. Arne Schröder* und *Dr. Jessica Steinbacher* zurück, die ein $^{15}\text{N}_2\text{-D}_2\text{-dC}$ in Zellverfütterungsexperimenten verwendeten. Da die verwendeten Verbindungen ein möglichst isotoopenreines Markierungsmuster aufweisen sollten, wurden auch andere Atompositionen für die Isotoopenmarkierung untersucht, da die Isotoopenreinheit des $^{15}\text{N}_2\text{-D}_2\text{-dC}$ -Sondenmoleküls stark durch die Isotoopenreinheit des kommerziell erhältlichen Ausgangsmaterials $^{15}\text{N}_2$ -Harnstoff (98% ^{15}N) begrenzt ist. Eine Weiterentwicklung des Isotoopenmarkierungsmusters führte zum $\text{D}_4\text{-dC}$ (synthetisiert im Rahmen meiner Master-Arbeit, betreut durch *Dr. René Rahimoff*) und schließlich zum $^{13}\text{C}_5\text{-D-dC}$ (entwickelt von *Dr. René Rahimoff*). Auch ein isotoopenmarkiertes Analogon des RNA-Bausteins Cytidin $^{13}\text{C}_5\text{-D-C}$ wurde von ihm synthetisiert. Abbildung 6-7 A zeigt einen Überblick der bisher entwickelten dC-basierten Sondenmoleküle und deren Isotoopenreinheit am Heterozyklus. Des Weiteren wird auch die Prozessierung des gefütterten $^{13}\text{C}_5\text{-D-dC}$ -Derivats im Genom der zu untersuchenden Zellen dargestellt (Abbildung 6-7 B). Dabei ist hervorzuheben, dass durch die selektive Isotoopenmarkierung der 5-Position das

verfütterte Sondenmolekül von seinem potentiellen Demethylierungsprodukt unterschieden werden kann.

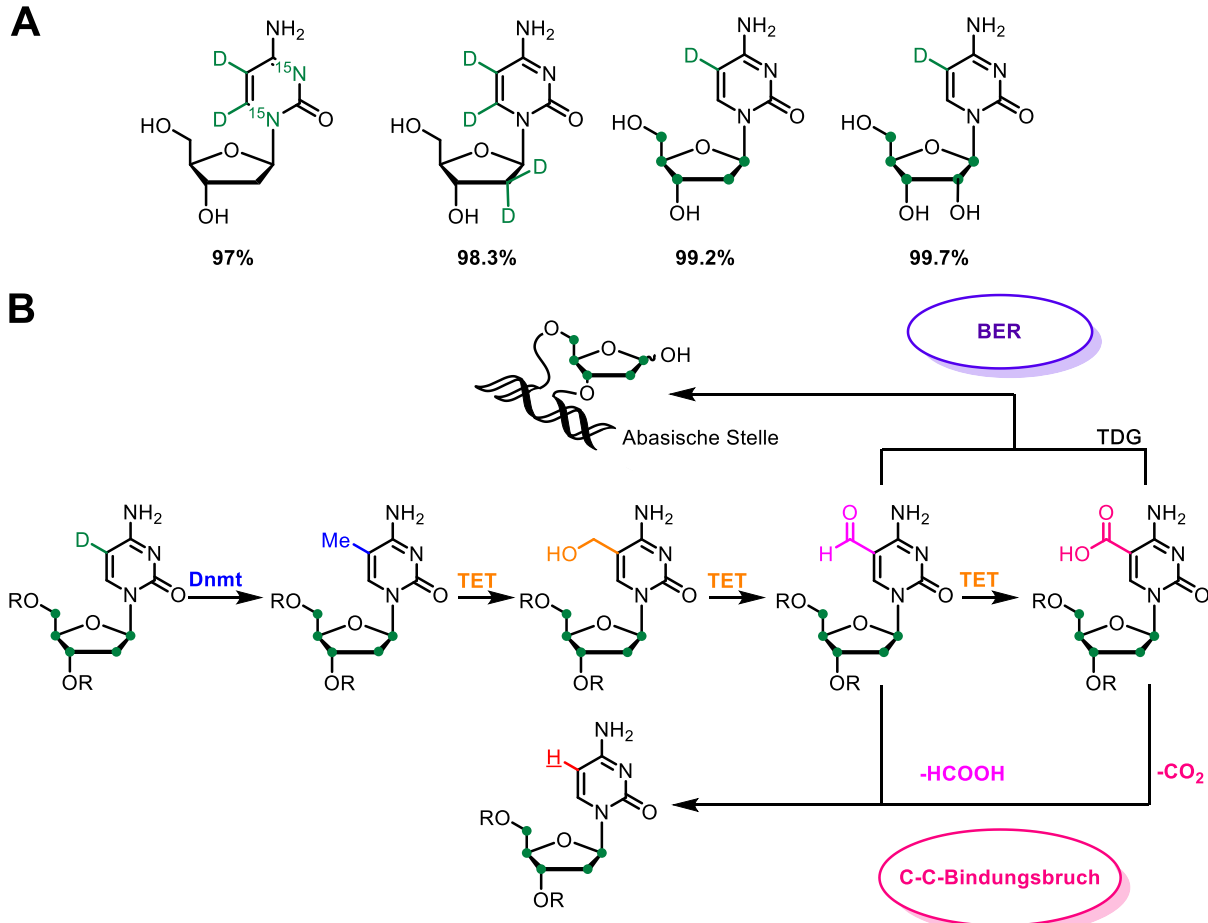


Abbildung 6-7: A) Isotopenmarkierte Verbindungen zur Untersuchung der direkten Deformylierung mit jeweiliger Isotopenreinheit; B) Zu erwartende biologische Prozessierung von $^{13}\text{C}_5\text{-D-dC}$. Die fetten Punkte symbolisieren das Isotop ^{13}C .

Auf Basis der Isotopenverfolgungsexperimente der von *Dr. René Rahimoff* synthetisierten $^{13}\text{C}_5\text{-D-dC}$ - und $^{13}\text{C}_5\text{-D-C}$ -Verbindungen, die von *Dr. Angie Kirchner* und *Dr. Katharina Iwan* durchgeführt wurden, konnten Hinweise gefunden werden, dass die gefütterten $^{13}\text{C}_5\text{-D}$ -Verbindungen *in vivo* in deren korrespondierende $^{13}\text{C}_5\text{-H}$ -Verbindungen umgewandelt, also eine direkte Demethylierung erfolgt sein könnte. Für einen umfassenden Überblick über die durchgeführten Experimente und die dabei erhaltenen Ergebnisse sei an dieser Stelle auf die Dissertationen von *Dr. René Rahimoff*, *Dr. Angie Kirchner* und *Dr. Katharina Iwan* verwiesen.^[114–116] Grundsätzlich kann aber festgestellt werden, dass eine Verunreinigung des Isotopenmarkierungsmusters des untersuchten $^{13}\text{C}_5\text{-D-dC}$ durch unerwünschtes

$^{13}\text{C}_5\text{-H-dC}$ (0.8%) die Datenauswertung erschwert. Das liegt daran, dass die Verunreinigung $^{13}\text{C}_5\text{-H-dC}$ mit dem potentiellen Deformylierungsprodukt von *in vivo* generiertem $^{13}\text{C}_5\text{-fdC}$ identisch ist. Die folgende Abbildung 6-8 zeigt die bei der Detektion $^{13}\text{C}_5\text{-D-dC}$ zu beobachtenden massenspektrometrischen Übergänge.

Ziel des hier dargestellten Projekts ist die Modifizierung der bereits bekannten Synthese von $^{13}\text{C}_5\text{-D-dC}$ mit dem Ziel eine noch höhere Isotopenreinheit als 99.2% im Zielmolekül zu erreichen. Damit soll ermöglicht werden, in anschließenden Zellverfütterungsexperimenten eindeutige Ergebnisse zu erhalten, die nicht durch eine mögliche $^{13}\text{C}_5\text{-H-dC}$ -Verunreinigung der verfütterten Verbindung herbeigeführt werden könnten.

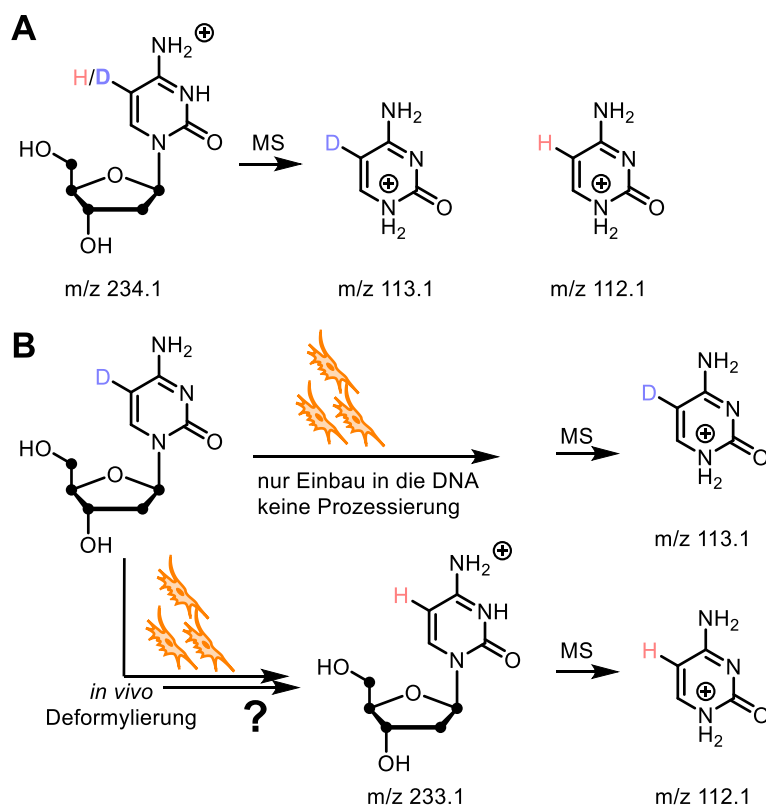


Abbildung 6-8: A) ESI-massenspektrometrische Fragmentierung von $^{13}\text{C}_5\text{-D-dC}$ sowie der nicht-deutierten Verunreinigung; B) Schema zur Veranschaulichung der ESI-massenspektrometrischen Fragmentierungen der Zielverbindung nach Zellverfütterung.

Das Einführen des Deuterium-Substituenten definiert demnach schlussendlich die Isotopenreinheit der $^{13}\text{C}_5\text{-D}$ -Spezies im Vergleich zur unerwünschten $^{13}\text{C}_5\text{-H}$ -Spezies. Da diese Reaktion allerdings in der letzten Stufe der Synthese durchgeführt wird, wurden die anfänglichen Syntheseschritte in Anlehnung an die bereits von *Dr. René Rahimoff* entwickelte Syntheseroute durchgeführt. Einzelne von mir entwickelte Modifikationen der bestehenden Vorschriften führten zu vergleichsweise leicht

verbesserten Ausbeuten bei der Nukleosidierung und auch bei der Einführung der *Markiewicz*-Schutzgruppe. Ein Überblick über die ersten Syntheseschritte ist in Abbildung 6-9 gegeben.

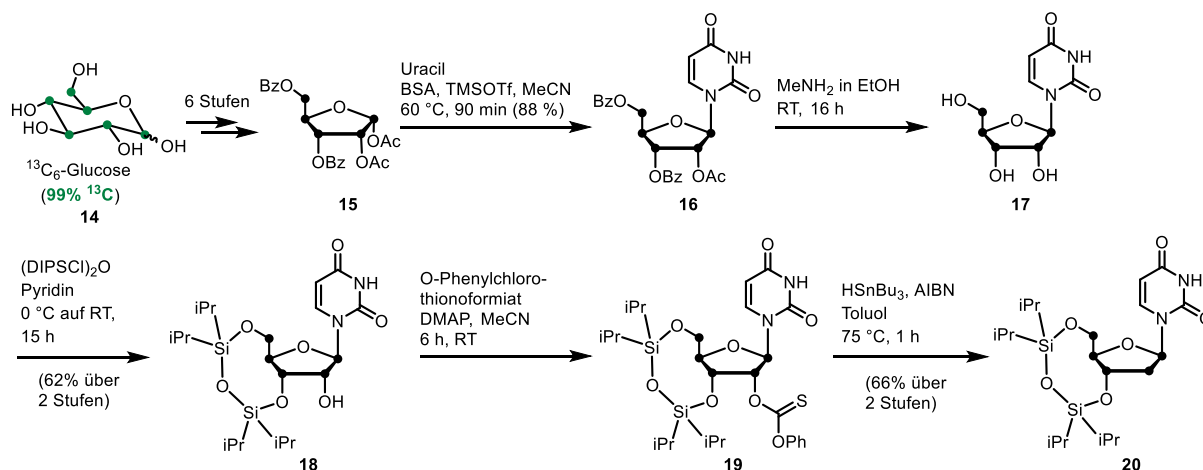


Abbildung 6-9: Syntheschema von geschütztem $^{13}\text{C}_5$ -dU **20** ausgehend von $^{13}\text{C}_6$ -Glucose **14**.

Die Quelle der ^{13}C -Markierung ist hierbei kommerziell erhältliche $^{13}\text{C}_6$ -Glukose **14**, deren Isotopenreinheit der Kohlenstoffmarkierungen bei 99 Atom-% ^{13}C liegt. Da in der endgültigen massenspektrometrischen Analyse der Massenübergang der Spaltung der glykosidischen Bindung untersucht wird, ist die demnach insbesondere die Isotopenreinheit des Pyrimidinheterozyklus entscheidend (siehe oben). Die Isotopenmarkierungen am Zucker sind hier nur notwendig, um die detektierten Moleküle vom biologischen Hintergrund der aus den Zellen isolierten, natürlichen Nukleoside zu unterscheiden. Dabei ist der natürliche Hintergrund fünffach isotopenmarkierter Nukleoside auf Grund des geringen, natürlichen Vorkommens an ^{13}C vernachlässigbar, weswegen die in der gegebenen Isotopenreinheit verwendete $^{13}\text{C}_6$ -Glukose **14** als geeignetes Startmaterial anzusehen ist.

Im Rahmen der Synthese wird $^{13}\text{C}_6$ -Glukose **14** über ein bekanntes sechsstufiges Verfahren in einen acylierten Ribose-Donor **15** für die anschließende Nukleosidierung überführt.^[51,59,95,117] Der Schlüsselschritt ist dabei eine Periodatspaltung, um die Hexose D-Glukose in D-Ribose, eine Pentose, umzuwandeln. Eine anschließende *Vorbrüggen*-Nukleosidierung für Ribonukleoside unter Verwendung des Ribose-Donors **15** und kommerziell erhältlichem Uracil lieferte basierend auf dem partizipierenden Effekt der Acetyl-Schutzgruppe am Sauerstoff der 2'-Position das isotopenmarkierte Uridin-Derivat **16** selektiv als β -Nukleosid. Eine darauffolgende Aminolyse der Acylschutzgruppen in ethanolischem Methylamin

resultierte im freien, isotopenmarkierten $^{13}\text{C}_5$ -Uridin **17**. Diesem wurde nach Isolation durch Gefriertrocknung Wasser entzogen, was bei der folgenden Einführung der *Markiewicz*-Schutzgruppe zu einer verbesserten Ausbeute von 62% über zwei Stufen verglichen mit den bisher bekannten Protokollen führte. Die auf diese Weise ungeschützte 2'-Hydroxylfunktion konnte in einem zweistufigen Verfahren zunächst in einen Thionester überführt und anschließend über eine *Barton-McCombie*-Deoxygenierung radikalisch entfernt werden, um so das geschützte $^{13}\text{C}_5$ -2'-Desoxyuridin-Derivat **20** zu erhalten.

In der bisher verwendeten Syntheseroute wurde dieses 2'-Desoxyuridin-Derivat **20** durch zweistufige Aminierung der 4-Position in die korrespondierende, isotopenmarkierte dC-Base umgewandelt und anschließend an der 5-Position oxidativ iodiert (siehe Abbildung 6-10, oben). Diese synthetischen Transformationen verlaufen allerdings über das geschützte $^{13}\text{C}_5$ -H-dC-Derivat **21**. Da dies eine potentielle Verunreinigung des Isotopenmusters durch Verschleppung dieser Verbindung über die nachfolgenden synthetischen Schritte zur Folge haben könnte, sollte dieser Schritt umgangen werden.

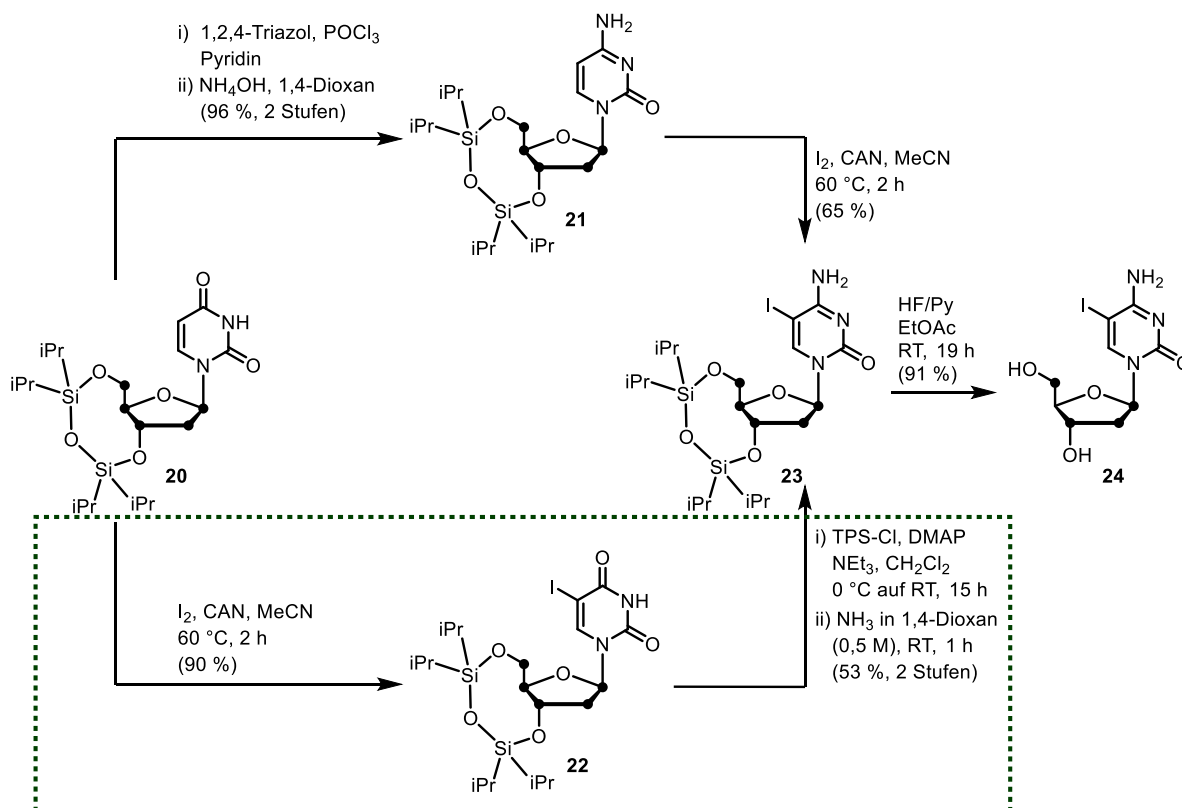
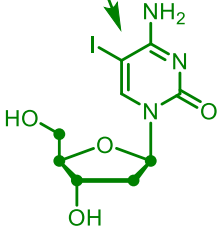


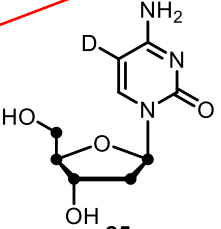
Abbildung 6-10: Schema der Iodierung, Aminierung und Entschützung zum isotopenmarkierten Iodo-dC **24**.


In der hier entwickelten, neuen Syntheseroute (siehe Abbildung 6-10, unten) wurde das geschützte $^{13}\text{C}_5$ -2'-Desoxyuridin-Derivat **20** zunächst oxidativ iodiert, um dann die iodierter Verbindung **22** durch Aminierung in dessen korrespondierende, iodierter dC-Base **23** umzuwandeln. Während die Iodierung unter analogen Bedingungen zur bekannten Syntheseroute durchgeführt werden konnte, musste für die Aminierung eine neue Methodologie entwickelt werden. Die erfolgreiche Aminierung gelang dabei in einem zweistufigen Verfahren, das zunächst die 4-Position durch portionsweise Zugabe von 2,4,6-Triisopropylbenzolsulfonylchlorid (TPS-Cl) in Gegenwart von Triethylamin und *Steglich*-Base in eine gute Abgangsgruppe überführte. Die Einführung der Aminofunktion wurde anschließend nach dünnschichtchromatographischem Nachweis des Verbrauchs des Startmaterials in einem Eintopfverfahren durch Zugabe von dioxanischem Ammoniak durchgeführt. Die direkte Aminierung im Eintopfverfahren umging dabei eine mögliche Hydrolyse des Sulfonsäureesters bei einer wässrigen Aufarbeitung zur Isolation des Zwischenproduktes oder bei dem Einsatz von wässrigem Ammoniak zur Aminierung. Eine potentielle Hydrolyse hätte dabei zur Rückreaktion zum 2'-Desoxyuridin-Derivat **22** führen und somit die Ausbeute der Reaktion mindern können. Zwar ist die Gesamtausbeute der neu entwickelten Synthesestrategie mit insgesamt 48% etwas niedriger als die der bisher verwendeten Route mit 62%, allerdings konnte das unerwünschte $^{13}\text{C}_5$ -H-dC-Derivat **21** erfolgreich synthetisch umgangen werden. Eine anschließende Entschützung von der Silylschutzgruppe mit Hydrogenfluorid in Pyridin (HF·Pyridin) lieferte in guter Ausbeute das freie $^{13}\text{C}_5$ -I-dC **24**, das nun in einem abschließenden Syntheseschritt an der 5-Position deuteriert werden sollte.

Die zur Deuterierung verwendete Methodik wurde bereits von *Dr. Arne Schröder* und *Dr. René Rahimoff* grundlegend entwickelt. Im Rahmen dieses Projekts sollte hier nun durch Modifizierungen der bestehenden Synthesevorschrift die Deuterierungseffizienz der Reaktion verbessert werden.

Abbildung 6-11 zeigt einen schematischen Überblick, welche Parameter der Reaktion beeinflusst wurden. Allerdings verschlechterte das Waschen des Palladium-Katalysators mit deuteriertem Aceton oder die Evaporation von NaHCO_3 aus einer Lösung in D_2O vor der Reaktion die erhaltene Ausbeute und Deuterierungseffizienz (<96%) sogar im Vergleich zu dem bekannten Verfahren. Auch die Verwendung einer Heliumatmosphäre vor Erzeugung der D_2 -Atmosphäre im Reaktionskolben

Gefriertrocknung

24

Waschen des Kat. mit Aceton-D₆
 Aus D₂O evaporieren
 Kat. mit D₂ beladen

25

Pd/C
 NaHCO₃, D₂
 D₃C-OD
 RT, über Nacht


$^{13}\text{C}_5\text{-D}/^{13}\text{C}_5\text{-H} > 99.9 \%$
 Ausbeute 64 %

Insgesamt wurde durch die so modifizierte Vorschrift $^{13}\text{C}_5\text{-D-dC}$ **25** mit einem Deuterierungsgrad von >99.9% am Heterozyklus in einer Ausbeute von 64% nach HPLC-Aufreinigung erhalten. Die hochaufgelösten MS/MS-Experimente zur Ermittlung des Deuterierungsgrades der synthetisierten Verbindung wurden von *Dr. Michael Stadlmeier* durchgeführt. Zusätzlich ist ein NMR-Spektrum der Verbindung in Abbildung 6-12 dargestellt. Obwohl die NMR-Analytik natürlich keine exakten Angaben über den Deuterierungsgrad der synthetisierten Verbindung geben kann, fällt hier dennoch interessanterweise ein Signal auf. Es handelt sich um das Protonensignal der 1'-Position, das neben der breiten Aufspaltung durch direkte Kopplung mit dem ^{13}C -Atom in 1'-Position auch ein kleines Signal in der Mitte der beiden Signale aufweist. Dieses Protonensignal resultiert aus einer Verunreinigung des Isotopenmusters des Startmaterials $^{13}\text{C}_6\text{-Glukose}$ **14**, dessen Isotopenreinheit mit 99 Atom-% ^{13}C angegeben wird. Im NMR-Spektrum kann hier die ^{12}C -Verunreinigung an der 1'-Position des synthetisierten Moleküls beobachtet werden. Leider können auf Grund der direkten ^1H - ^{13}C -Kopplungen und des daraus resultierenden relativ unübersichtlichen Spektrums keine weiteren ^{12}C -Verunreinigungen detektiert werden. Des Weiteren fällt auf, dass im hier

gemessenen Spektrum, wie zu erwarten, kein Signal der deuterierten 5-Position zu sehen ist, was auch für eine sehr isotopenreine Zielverbindung spricht.

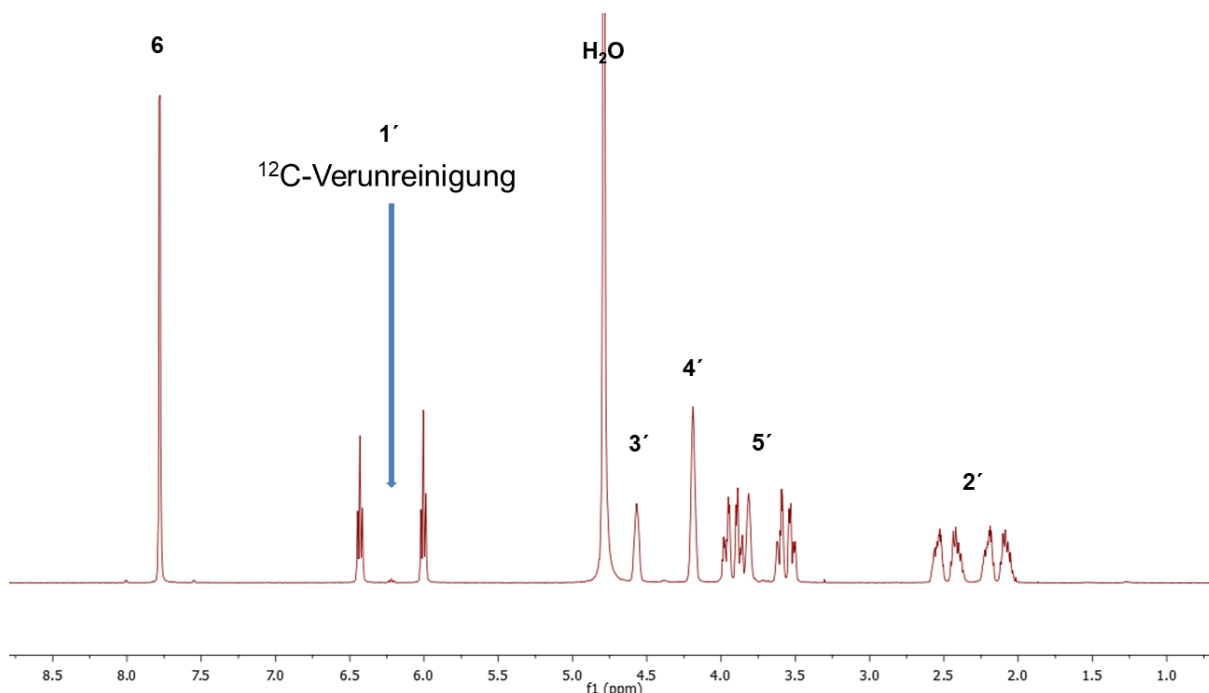


Abbildung 6-12: NMR-Spektrum der Zielverbindung **25** und Darstellung der ^{12}C -Verunreinigung.

Nach der Synthese der Zielverbindung **25**, die basierend auf der neu entwickelten Syntheseroute nun einen nahezu vollständigen Deuterierungsgrad am Heterozyklus aufweist, soll diese in Verfütterungsexperimenten eingesetzt werden. Im Folgenden soll nur ein kurzer Überblick über das durchgeführte Experiment und dessen Ergebnisse gegeben werden, für detaillierte Beschreibungen der Zellkultur-Experimente und der massenspektrometrischen Datenanalyse sei auf die zukünftigen Dissertationen von *Ewelina Kaminska* und *Florian Schelter* verwiesen.

Bei dem zellbiologisch durchgeführten Verfütterungsexperiment handelte es sich um ein sogenanntes *Pulse-Chase*-Experiment (siehe Abbildung 6-13). Dafür werden humane, induzierte pluripotente Stammzellen (iPSCs) verwendet, deren Neurogenese durch Doxycyclin-Zugabe induziert werden kann (*small molecule-inducible neurogenin iPS-line* (iNGNs)). Untersuchungen von *Dr. Jessica Steinbacher* und *Dr. Angie Kirchner* bestätigten, dass die Wachstumsrate der Zellen und somit die DNA-Neusynthese nach Induktion der Differenzierung stark abnimmt. Während innerhalb der ersten 24 h nach Induktion noch reduziert stattfindende DNA-Neusynthese beobachtet werden kann, ist diese nach 48 h und später nicht mehr detektierbar. Nach Kultivierung der Zellen für 48 h wird dem Medium $^{13}\text{C}_5\text{-D-dC}$ und

zwölffach isotopenmarkiertes dT (dT^{+12}) für 24 h zugesetzt (*Pulse*). Anschließend wird das Medium verworfen, neues Medium ohne die isotopenmarkierten Verbindungen zugesetzt und die Neurogenese durch Doxycyclin-Zugabe induziert. Zellproben zur Untersuchung der isotopenmarkierten Nukleoside in der DNA werden vor der Induktion und anschließend alle 24 h gesammelt (*Chase*).

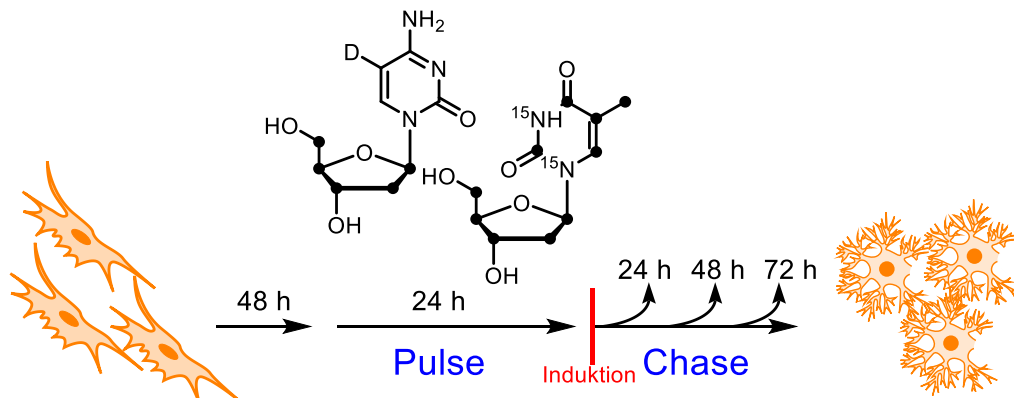


Abbildung 6-13: Überblick über das durchgeführte *pulse-chase*-Experiment.

Da das isotopenmarkierte Material nur während des *Pulse* verfüttert wurde und die DNA-Neusynthese spätestens 48 h nach Induktion zum Erliegen kommt, sollten bei diesem Experiment die Prozesse an den isotopenmarkierten Nukleosiden im Genom der Zellen untersucht werden können. Über die Bestimmung des Gehalts an dT^{+12} in der DNA kann letztendlich bei der massenspektrometrischen Analyse auf das Auftreten von noch aktiver DNA-Neusynthese nach der Induktion rückgeschlossen werden. Denn erst nach dem Erliegen der DNA-Neusynthese können aussagekräftige Ergebnisse zur aktiven Demethylierung erhalten werden, da potentielle, eine Masseneinheit leichtere Demethylierungsprodukte nicht in die DNA eingebaut werden können.

Bei der Analyse der massenspektrometrischen Daten konnte eine sehr geringe Inkorporationsrate von ca. 2% von $^{13}\text{C}_5\text{-D-dC}$ in die DNA der Zellen detektiert werden. Diese niedrige Inkorporationsrate wurde allerdings auch schon bei den Experimenten von *Dr. Angie Kirchner* festgestellt und kann vermutlich auf starke Desaminierung der verfütterten Verbindung zurückgeführt werden.

Auf Grund der geringen Abundanz der isotopenmarkierten $^{13}\text{C}_5\text{-D-dC}$ -Verbindung im Genom konnte nur dessen Methylierungsprodukt $^{13}\text{C}_5\text{-mdC}$ quantifiziert werden. Ein Signal von $^{13}\text{C}_5\text{-hmdC}$ konnte zwar detektiert, auf Grund der geringen Intensität allerdings nicht quantifiziert werden. Die noch selteneren, höher oxidierten Derivate

$^{13}\text{C}_5\text{-fdC}$ und $^{13}\text{C}_5\text{-cadC}$ sowie ein potentielles Deformylierungsprodukt der verfütterten Verbindung konnten bei der massenspektrometrischen Analyse nicht gefunden werden. Ein Überblick ist in Abbildung 6-14 gegeben.

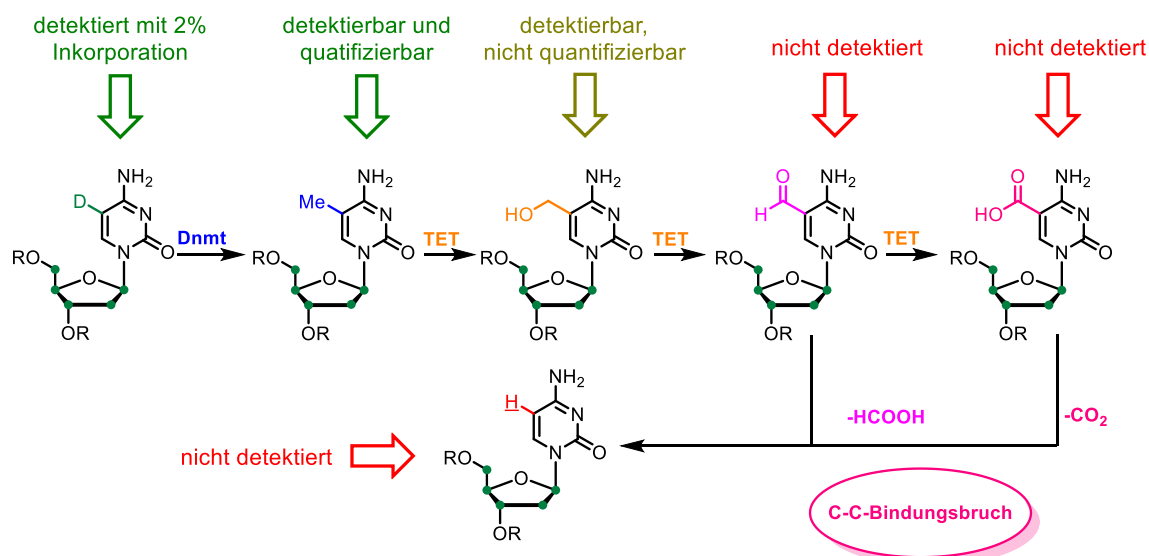


Abbildung 6-14: Erste Ergebnisse der Verfütterungsexperimente von $^{13}\text{C}_5\text{-D-dC}$.

6.2.2. Chemische Derivatisierung zur Verbesserung der Signalintensität

Da die massenspektrometrische Analytik wie oben beschrieben basierend auf der geringen Inkorporation der Verbindung **25** in die DNA sowie ebenfalls basierend auf der niedrigen Ausbeute an isolierter DNA nicht möglich war, wurde im Folgenden versucht durch chemische Derivatisierung die Signalintensitäten von $^{13}\text{C}_5\text{-fdC}$ und $^{13}\text{C}_5\text{-cadC}$ zu verstärken. Kürzlich wurde eine Reihe neuer, Hydrazin-basierter Derivatisierungsreagenzien vorgestellt, deren Addukte mit fdC und cadC zu einer deutlichen Intensitätssteigerung der Signale im Vergleich mit den unreaktierten Nukleosiden bei TQ-massenspektrometrischen Messungen aufwiesen.^[118,119] Da die chemische Derivatisierung gemäß der Literatur direkt nach dem enzymatischen Verdau der DNA in einem einfachen, zweistufigen Verfahren durchführbar sei,^[119] wurde untersucht, ob diese Methodik im Rahmen dieses Projekts angewandt werden könnte. Das effektivste der in der Literatur beschriebenen Derivatisierungsreagenzien wurde gemäß der Vorschriften von *Zhao et al.*^[118] in einem zweistufigen Verfahren ausgehend von Cyanurchlorid **26** synthetisiert (siehe Abbildung 6-15).

Anschließend wurden im präparativen Maßstab die Derivatisierungsprodukte des Reagenzes mit fdC und cadC hergestellt, wobei für diese Synthesen die

Literaturvorschriften zur Reaktion der Nukleoside mit **27** modifiziert werden mussten. Die Reaktion von **27** mit fdC wurde dazu in einer Lösung von 10% Eisessig in Methanol bei 37 °C für 4 h und die Derivatisierung von cadC mit **27** unter Aktivierung durch HOBt und EDC-HCl in Acetonitril bei 37 °C für 3 h durchgeführt. Insgesamt konnten nach HPLC-Aufreinigung die Reaktionsprodukte von **27** mit fdC und cadC im präparativen Maßstab in mäßiger bis schlechter Ausbeute erhalten werden. Auch das Reaktionsprodukt von 2'-Desoxy-D-ribose mit **27** wurde in analogen Bedingungen zur Reaktion von **27** mit fdC synthetisiert. Dieses sollte ebenfalls zur massenspektrometrischen Methodenentwicklung verwendet werden, da angenommen wurde, dass das Derivatisierungsreagenz **27** unter den verwendeten Reaktionsbedingungen auch mit potentiell in der DNA vorhandenen abasischen Stellen reagieren könnte.

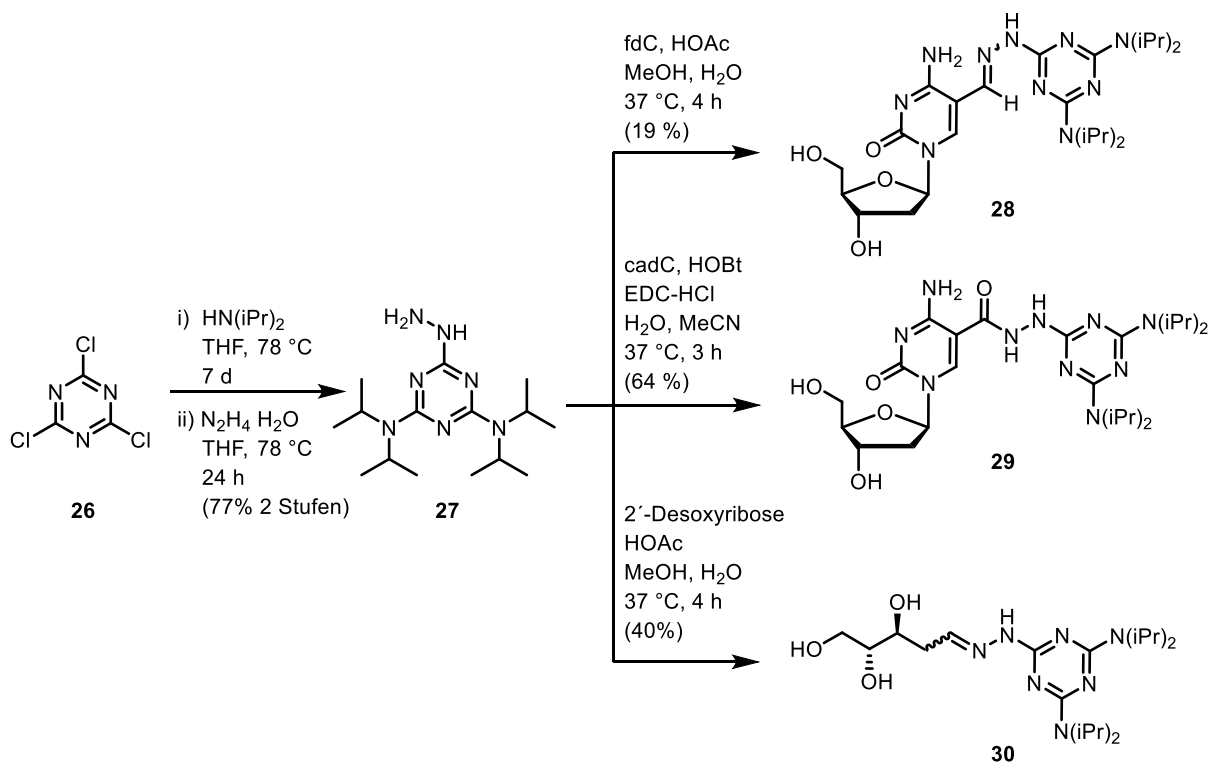


Abbildung 6-15: Schema der Synthese des Derivatisierungsreagenzes und der Derivatisierungen von fdC, cadC und 2'-Desoxyribose.

Durch Derivatisierung der isotopenmarkierten abasischen Stellen und anschließender massenspektrometrischer Detektion könnten die mit der direkten Demethylierung konkurrierenden BER-Prozesse gleichzeitig quantifiziert werden. Die Methoden zur massenspektrometrischen Detektion wurden von *Florian Schelter* entwickelt. Die Durchführung der Derivatisierungsreaktionen mit **27** nach dem Verdau

der DNA-Proben wurden ebenfalls von ihm nach den bekannten Protokollen durchgeführt.^[119]

Grundsätzlich konnte bei den Untersuchungen festgestellt werden, dass die chemische Derivatisierung von fdC, cadC und abasischen Stellen mit dem Derivatisierungsreagenz **27** zu einer Signalverstärkung bei der Detektion der resultierenden Derivate bei der ESI-MS/MS-Analytik wie in der Literatur beschrieben führte. Da im Rahmen dieser Messungen größere Mengen an DNA verglichen mit den oben beschriebenen Experimenten enzymatisch verdaut und massenspektrometrisch analysiert wurden, konnte nun auch $^{13}\text{C}_5\text{-hmdC}$ in den Proben detektiert und quantifiziert werden. Die Detektion des hmdC-Derivats der verfütterten Verbindung **25** konnte parallel zur Derivatisierung und Untersuchung der höher oxidierten Modifikationen durchgeführt werden. Die Menge an $^{13}\text{C}_5\text{-hmdC}$ in der untersuchten DNA war in den untersuchten Proben im Verlauf der verschiedenen Tage des *Pulse-Chase*-Experiments durchgehend konstant und konnte mit Hilfe eines bestehenden isotoopenmarkierten $\text{D}_2\text{-hmdC}$ -Standards quantifiziert werden (siehe Abbildung 6-16).

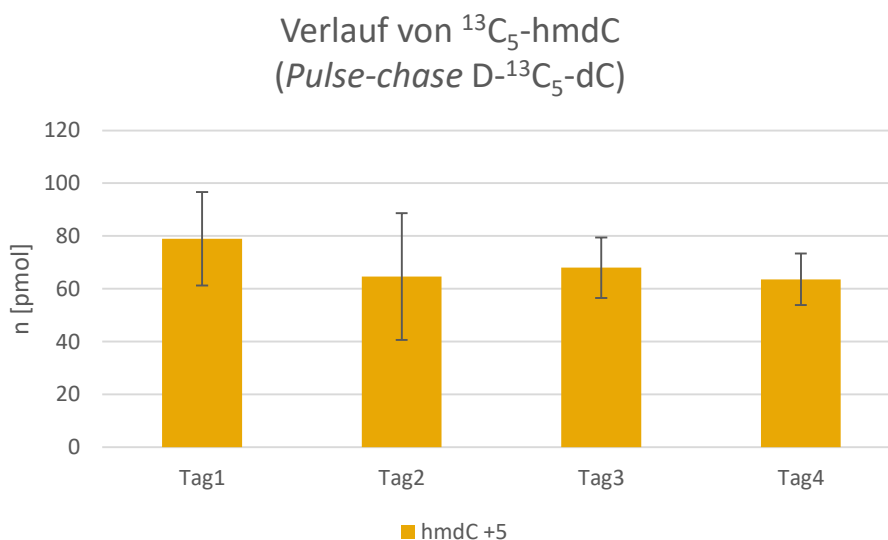


Abbildung 6-16: Massenspektrometrische Quantifizierung von $^{13}\text{C}_5\text{-hmdC}$ im Verlauf des Experiments.

Wie bereits erwähnt konnten auch die Derivate von $^{13}\text{C}_5\text{-fdC}$ und $^{13}\text{C}_5\text{-cadC}$ nach Reaktion mit dem Hydrazinreagenz **27** detektiert werden. Da zur exakten Quantifizierung weder ein isotoopenmarkierter Standard der derivatisierten Verbindungen zur Verfügung stand noch exakte Kalibriergeraden gemessen worden waren, wurde eine näherungsweise Quantifizierung basierend auf zugegebenen

isotopenmarkierten Isotopologen von fdC und cadC durchgeführt (siehe Abbildung 6-17).

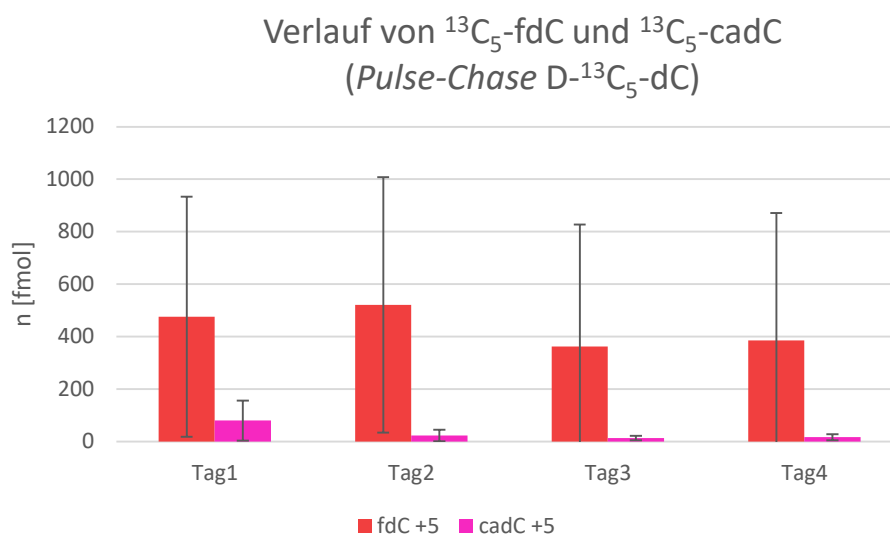


Abbildung 6-17: Massenspektrometrische Quantifizierung von $^{13}\text{C}_5$ -fdC und $^{13}\text{C}_5$ -cadC im Verlauf des Experiments.

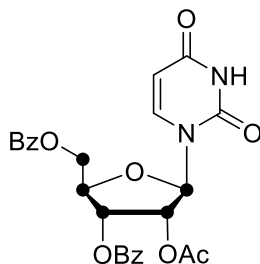
Dabei wurde von einer quantitativen Umsetzung von $^{13}\text{C}_5$ -fdC und $^{13}\text{C}_5$ -cadC sowie der dem Verdauungsmisch zugegebenen Verbindungen $^{15}\text{N}_2$ -fdC und $^{15}\text{N}_2$ -cadC mit dem Derivatisierungsreagenz **27** ausgegangen. Diese Annahme basiert auf den Studien von Yu *et al.*,^[119] die unter den eingesetzten Derivatisierungsbedingungen quantitativen Umsatz von cadC und fdC beobachten konnten. Basierend auf diesen Annahmen wurden durch direkten Vergleich der erhaltenen Signale bei der massenspektrometrischen Analytik erste genäherte Quantifizierungsergebnisse für $^{13}\text{C}_5$ -fdC und $^{13}\text{C}_5$ -cadC ermittelt. Dabei liegen die Werte von $^{13}\text{C}_5$ -cadC jeweils eine Größenordnung niedriger als die ermittelten Werte für $^{13}\text{C}_5$ -fdC, was im biologischen Kontext auf Basis der natürlichen Vorkommensraten von fdC und cadC sinnvoll erscheint. Da allerdings die Quantifizierung nicht exakt über die Methode der isotopenmarkierten Standards oder über Kalibriergeraden durchgeführt wurde,^[120] mussten hier große Abweichungen bei den Bestimmungen von $^{13}\text{C}_5$ -fdC und $^{13}\text{C}_5$ -cadC innerhalb der verschiedenen Messungen hingenommen werden. Eine Wiederholung der Experimente unter Einsatz exakter Quantifizierungsmethoden ist für die endgültige Bestimmung der höher oxidierten, isotopenmarkierten Modifikationen notwendig.

6.2.3. Projektbeiträge

Im Rahmen dieses Projekts synthetisierte ich das isotopenmarkierte $^{13}\text{C}_5\text{-D-dC}$, das für die *Pulse-Chase*-Experimente in der Zellkultur verwendet wurde. Die Synthesevorschriften beruhen dabei auf Arbeiten von *Dr. René Rahimoff* und *Dr. Arne Schröder*, die allerdings an einigen Stellen modifiziert wurden. Außerdem synthetisierte ich das Hydrazin-basierte Derivatisierungsreagenz und stellte dessen Addukte mit fdC, cadC sowie 2'-Desoxyribose her, die dann für die Entwicklung massenspektrometrischer Methoden von *Florian Schelter* verwendet wurden. Alle Arbeiten in der Zellkultur wurden von *Ewelina Kaminska* durchgeführt. DNA-Verdaus, Derivatisierungen der verdauten Nukleosidmischungen und hochsensitive massenspektrometrische Messungen wurden von *Florian Schelter* und *Eva Korytiaková* durchgeführt. *Dr. Michael Stadlmeier* führte hochaufgelöste MS/MS-Messungen zur Bestimmung der Isotopenreinheit der von mir synthetisierten $^{13}\text{C}_5\text{-D-dC}$ -Verbindung durch.

6.2.4. Experimentalteil

2'-O-Acetyl-3',5'-di-O-benzoyl-(1',2',3',4',5'-¹³C₅)-uridin (16)



In einem aufgeheizten Schlenk-Kolben wurden 310 mg β -1',2'-Di-O-Acetyl-3',5'-di-O-benzoyl-(1',2',3',4',5'-¹³C₅)-ribofuranose^[51,59,95,117] **15** (0.69 mmol, 1.0 eq.) und 85 mg Uracil (0.76 mmol, 1.1 eq.) in 3.9 mL trockenem MeCN unter Argon-Atmosphäre suspendiert. Der Suspension wurden 0.51 mL *N,O*-Bis(trimethylsilyl)acetamid (BSA, 2.08 mmol, 3.0 eq.) zugegeben und das Gemisch wurde 30 min bei 60 °C verrührt. Zu der erhaltenen Lösung wurden dann 0.44 mL Trimethylsilyltrifluormethansulfonat (TMSOTf, 2.43 mmol, 3.5 eq.) zugetropft und das Gemisch wurde weitere 50 min bei 60 °C gerührt. Daraufhin wurde das Reaktionsgemisch in 200 mL gesättigte wässrige NaHCO₃-Lösung gegossen und mit CH₂Cl₂ (3x 200 mL) extrahiert. Die vereinigten organischen Phasen wurden über Na₂SO₄ getrocknet und Lösungsmittel wurden *in vacuo* entfernt. Das Rohprodukt wurde mittels Säulenchromatographie an Silicagel (iHex:EtOAc 1:1 → iHex:EtOAc 1:2) aufgereinigt, um 303 mg des Nukleosidierungsproduktes **16** (0.61 mmol, 88 %) als gelblichen Schaum zu erhalten.

¹H-NMR (400 MHz, CDCl₃): δ /ppm = 8.34 (s, 1H, N-H), 8.10 – 8.02 (m, 4H, Bz-H), 7.67 – 7.57 (m, 2H, Bz-H), 7.52 – 7.45 (m, 4H, Bz-H), 7.34 (dd, J = 8.2, 2.7 Hz, 1H, 6-H), 6.20 (ddd, J = 169.1, 5.8, 3.2 Hz, 1H, 1'-H), 5.75 (dt, J = 158.2, 3.6 Hz, 1H, 3'-H), 5.57 (dd, J = 8.2, 2.3 Hz, 1H, 5-H), 5.54 (d, $^1J_{C-H}$ = 152.7 Hz, 1H, 2'-H), 4.79 (d, $^1J_{C-H}$ = 150.6 Hz, 1H, 5'-H), 4.64 (d, $^1J_{C-H}$ = 152.5 Hz, 1H, 4'-H), 4.61 (d, $^1J_{C-H}$ = 149.8 Hz, 1H, 5'-H), 2.06 (s, 3H, H₃C-C=O).

¹³C-NMR (101 MHz, CDCl₃): δ /ppm = 169.83 (H₃C-C=O), 166.12 (Ph-C=O), 165.47 (Ph-C=O), 162.29 (C-4), 150.05 (C-2), 139.27 (C-6), 134.05 (2 x C_{Ar}), 133.89 (2 x C_{Ar}), 129.98 (2 x C_{Ar}), 129.75 (2 x C_{Ar}), 128.95 (2 x C_{Ar}), 128.84 (2 x C_{Ar}), 103.60 (C-5), 87.40 (dd, J = 44.5, 3.7 Hz, C-1'), 80.56 (dd, J = 43.0, 38.2 Hz, C-4'), 73.10 (dd, J

= 44.4, 39.9 Hz, C-2'), 71.08 (td, $J = 40.0, 3.8$ Hz, C-3'), 63.81 (d, $J = 43.0$ Hz, C-5'), 20.57 (H₃C-C=O).

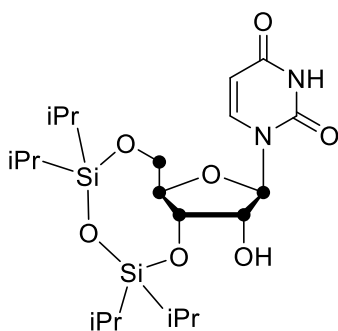
HRMS (ESI⁺): m/z berechnet für C₂₀¹³C₅H₂₆N₃O₉⁺ [M+NH₄]⁺ 517.1831; gefunden: 517.1832.

HRMS (ESI⁻): m/z berechnet für C₂₀¹³C₅H₂₁N₂O₉⁻ [M-H]⁻ 498.1420; gefunden: 498.1420.

IR (ATR): ν (cm⁻¹) = 3069 (w), 1719 (s), 1695 (s), 1452 (m), 1375 (w), 1265 (s), 1225 (m), 1109 (m), 1070 (m), 1026 (w), 710 (s).

R_f (iHex:EtOAc 1:1): 0.21.

3',5'-(1,1,3,3-Tetraisopropyldisiloxan-1,3-diyl)-(1',2',3',4',5'-¹³C₅)-uridin (18)



300 mg des geschützten Uridin-Derivats **16** (0.60 mmol, 1.0 eq.) wurden bei Raumtemperatur mit 4.8 mL ethanolischem MeNH₂ versetzt (33 Gew.-%) und 16 h verrührt. Das erhaltene Reaktionsgemisch wurde *in vacuo* zur Trockene eingeeengt und der erhaltene Rückstand wurde in 75 mL ddH₂O aufgenommen. Die wässrige Phase wurde anschließend mit CH₂Cl₂ (5 x 75 mL) extrahiert, danach *in vacuo* auf ca. 15 mL reduziert und schließlich lyophilisiert, um das freie Nukleosid **17** zu liefern, das ohne weitere Aufreinigung im nächsten Schritt eingesetzt wurde.

Die entschützte Verbindung **17** wurde in 4.4 mL trockenem Pyridin gelöst und unter Argonatmosphäre auf 0 °C gekühlt. Der Lösung wurden bei 0 °C 0.22 mL 1,3-Dichlor-1,1,3,3-tetraisopropyldisiloxan (0.69 mmol, 1.15 eq.) tropfenweise zugegeben, das Reaktionsgemisch wurde auf Raumtemperatur erwärmen gelassen und 15 h verrührt. Das Lösungsmittel wurde *in vacuo* entfernt und der Rückstand wurde mit Toluol coevaporiert (1 x 15 mL). Anschließend wurde dieser in 100 mL CH₂Cl₂ aufgenommen und mit gesättigter wässriger NH₄Cl-Lösung (100 mL)

gewaschen. Die wässrige Phase wurde daraufhin noch weiter mit CH_2Cl_2 (2 x 100 mL) extrahiert, die vereinigten organischen Phasen wurden über Na_2SO_4 getrocknet und Lösungsmittel wurden *in vacuo* entfernt. Säulenchromatographie an Silicagel (CH_2Cl_2 :MeOH 50:1) zur Aufreinigung des Rohprodukts lieferte 180 mg der silylierten Verbindung **18** (0.37 mmol, 62 % über 2 Stufen) als farblosen Schaum.

$^1\text{H-NMR}$ (600 MHz, CDCl_3): δ/ppm = 8.15 (s, 1H, N-H), 7.66 (dd, J = 8.2, 2.6 Hz, 1H, 6-H), 5.72 (d, $^1J_{\text{C-H}}$ = 174.4 Hz, 1H, 1'-H), 5.69 (dd, J = 8.1, 2.1 Hz, 1H, 5-H), 4.54 – 3.85 (m, 5H, 2'-H, 3'-H, 4'-H, 5'-H), 1.11 – 1.00 (m, 28H, Si-CH, Si-CH(CH_3)₂).

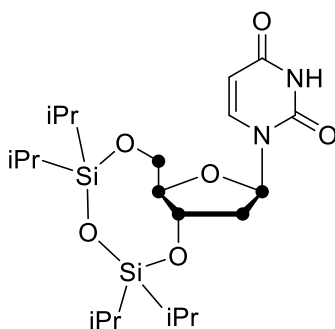
$^{13}\text{C-NMR}$ (151 MHz, CDCl_3): δ/ppm = 162.81 (C-4), 149.75 (C-2), 140.21 (C-6), 102.02 (C-5), 91.01 (dt, J = 43.2, 2.8 Hz, C-1'), 82.11 (t, J = 42.3 Hz, C-4'), 75.34 (dd, J = 43.4, 37.8 Hz, C-2'), 69.27 (ddd, J = 41.2, 37.9, 3.1 Hz, C-3'), 60.50 (d, J = 43.4 Hz, C-5'), 17.59 (Si-CH(CH_3)₂), 17.51 (Si-CH(CH_3)₂), 17.42 (Si-CH(CH_3)₂), 17.39 (Si-CH(CH_3)₂), 17.18 (Si-CH(CH_3)₂), 17.11 (Si-CH(CH_3)₂), 17.06 (Si-CH(CH_3)₂), 16.98 (Si-CH(CH_3)₂), 13.52 (Si-CH(CH_3)₂), 13.08 (Si-CH(CH_3)₂), 13.04 (Si-CH(CH_3)₂), 12.67 (Si-CH(CH_3)₂).

HRMS (ESI⁺): m/z berechnet für $\text{C}_{16}^{13}\text{C}_5\text{H}_{39}\text{N}_2\text{O}_7\text{Si}_2^+$ $[\text{M}+\text{H}]^+$ 492.2458; gefunden: 492.2462, m/z berechnet für $\text{C}_{16}^{13}\text{C}_5\text{H}_{41}\text{N}_2\text{O}_8\text{Si}_2^+$ $[\text{M}+\text{H}_2\text{O}+\text{H}]^+$ 510.2564; gefunden: 510.2567.

IR (ATR): ν (cm^{-1}) = 3355 (br w), 2945 (m), 2867 (m), 1683 (s), 1463 (m), 1387 (w), 1261 (m), 1084 (s), 1035 (s), 920 (w), 884 (s), 857 (m), 808 (m), 693 (m).

R_f (CH_2Cl_2 :MeOH 40:1): 0.26.

3',5'-(1,1,3,3,-Tetraisopropyldisiloxan-1,3-diyl)-2'-desoxy-(1',2',3',4',5'- $^{13}\text{C}_5$)-uridin (20)



173 mg des Uridin-Derivats **18** (0.35 mmol, 1.0 eq.) wurden in 4.9 mL trockenem MeCN gelöst und auf 0 °C gekühlt. Anschließend wurden der Lösung 279 mg DMAP (2.29 mmol, 6.5 eq.) und 58 µL O-Phenyl-chlorothionoformiat (0.42 mmol, 1.2 eq.) zugesetzt. Das Reaktionsgemisch wurde daraufhin auf Raumtemperatur erwärmen gelassen und 6 h verrührt. Zu der erhaltenen gelben Lösung wurden 5.0 mL trockenes MeOH zugegeben und das Gemisch wurde *in vacuo* zur Trockene eingeeengt. Der Rückstand wurde dann in 50 mL CH₂Cl₂ aufgenommen und mit gesättigter wässriger NaHCO₃-Lösung (50 mL) gewaschen. Die wässrige Phase wurde dann weiter mit CH₂Cl₂ (2 x 50 mL) extrahiert, die vereinigten organischen Phasen über Na₂SO₄ getrocknet und Lösungsmittel wurden *in vacuo* entfernt. Das Rohprodukt wurde mittels Säulenchromatographie an Silicagel (iHex:EtOAc 3:1 → iHex:EtOAc 1:1) aufgereinigt, um 145 mg des Produkts **19** (0.23 mmol) als weißen Schaum zu liefern.

Unter N₂-Atmosphäre wurden 145 mg des Zwischenprodukts **19** (0.23 mmol, 1.0 eq.) in 4.6 mL trockenem Toluol gelöst. Anschließend wurden 8 mg AIBN (46 µmol, 0.2 eq.) und 92 µL HSnBu₃ (0.35 mmol, 1.5 eq.) hinzugegeben und das Gemisch wurde gründlich entgast (*freeze, pump, thaw* 3x). Dann wurde die Lösung für 1 h auf 75 °C erwärmt. Weitere 30 µL HSnBu₃ (0.12 mmol 0.5 eq.) wurden zugetropft und das Gemisch wurde für 1 h weiter bei 75 °C verrührt. Nach vollständigem Umsatz des Startmaterials wurde das Reaktionsgemisch auf Raumtemperatur abkühlen gelassen und *in vacuo* eingeeengt. Der Rückstand wurde in CH₂Cl₂ (50 mL) aufgenommen und mit H₂O (50 mL) gewaschen. Die wässrige Phase wurde weiter mit CH₂Cl₂ (2 x 50 mL) extrahiert, die vereinigten organischen Phasen wurden über Na₂SO₄ getrocknet, und *in vacuo* zur Trockene eingeeengt. Das Rohprodukt wurde mittels Säulenchromatographie an Silicagel (iHex:EtOAc 3:1 → iHex:EtOAc 2:1) aufgereinigt, um 110 mg des Deoxygenierungsproduktes **20** (0.23 mmol, 66 % über 2 Stufen) als farblosen Schaum zu erhalten.

¹H-NMR (400 MHz, CDCl₃): δ/ppm = 8.04 (s, 1H, N-H), 7.77 (dd, *J* = 8.1, 2.5 Hz, 1H, 6-H), 6.03 (dd, *J* = 175.1, 7.0 Hz, 1H, 1'-H), 5.71 – 5.64 (d, ¹*J*_{C-H} = 8.5 Hz, 1H, 5-H), 4.43 (d, ¹*J*_{C-H} = 144.0 Hz, 1H, 3'-H), 4.41 – 3.51 (m, 3H, 4'-H, 5'-H), 2.51 (d, ¹*J*_{C-H} = 133.5 Hz, 1H, 2'-H), 2.27 (d, ¹*J*_{C-H} = 136.3 Hz, 1H, 2'-H), 1.13 – 0.99 (m, 28H, Si-CH, Si-CH(CH₃)₂).

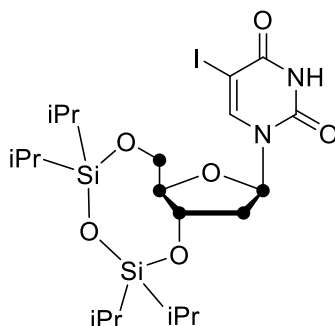
^{13}C -NMR (101 MHz, CDCl_3): δ/ppm = 162.87 (C-4), 149.86 (C-2), 139.83 (C-6), 101.78 (C-5), 85.24 (ddd, J = 43.1, 40.6, 2.4 Hz, C-4'), 84.50 (dt, J = 36.3, 2.1 Hz, C-1'), 67.12 (t, J = 39.6 Hz, C-3'), 60.11 (d, J = 43.3 Hz, C-5'), 40.04 (ddt, J = 38.5, 36.5, 2.4 Hz, C-2'), 17.65 (Si-CH(CH_3)₂), 17.57 (Si-CH(CH_3)₂), 17.48 (Si-CH(CH_3)₂), 17.42 (Si-CH(CH_3)₂), 17.21 (Si-CH(CH_3)₂), 17.10 (2 x Si-CH(CH_3)₂), 16.97 (Si-CH(CH_3)₂), 13.59 (Si-CH(CH_3)₂), 13.17 (Si-CH(CH_3)₂), 13.12 (Si-CH(CH_3)₂), 12.57 (Si-CH(CH_3)₂).

HRMS (ESI⁺): m/z berechnet für $\text{C}_{16}^{13}\text{C}_5\text{H}_{39}\text{N}_2\text{O}_6\text{Si}_2^+$ $[\text{M}+\text{H}]^+$ 476.2509; gefunden: 476.2513, m/z berechnet für $\text{C}_{16}^{13}\text{C}_5\text{H}_{41}\text{N}_2\text{O}_7\text{Si}_2^+$ $[\text{M}+\text{H}_2\text{O}+\text{H}]^+$ 494.2615; gefunden: 494.2618.

IR (ATR): ν (cm^{-1}) = 3410 (br w), 2944 (s), 2897 (s), 1682 (s), 1463 (m), 1386 (w), 1263 (m), 1083 (s), 1054 (s), 884 (s), 858 (m), 690 (m).

R_f (iHex:EtOAc 2:1): 0.29.

3',5'-(1,1,3,3,-Tetraisopropyldisiloxan-1,3-diyl)-5-iodo-2'-desoxy-(1',2',3',4',5'- $^{13}\text{C}_5$)-uridin (22**)**



110 mg der Verbindung **20** (0.23 mmol, 1.0 eq.) wurden in 13.8 mL trockenem MeCN gelöst, mit 117 mg elementarem Iod (0.46 mmol, 2.0 eq.) und 254 mg Cer(IV)-ammoniumnitrat (CAN, 0.46 mmol, 2.0 eq.) versetzt und das Gemisch wurde 2 h auf 60 °C erwärmt. Dann wurden 30 mL gesättigte wässrige NaHCO_3 -Lösung und 30 mL gesättigte wässrige $\text{Na}_2\text{S}_2\text{O}_3$ -Lösung zugegeben und MeCN wurde *in vacuo* entfernt. Daraufhin wurde mit CH_2Cl_2 (3 x 100 mL) extrahiert, vereinigte organische Phasen wurden über Na_2SO_4 getrocknet und Lösungsmittel wurden *in vacuo* entfernt. Das Rohprodukt wurde über Säulenchromatographie an Silicagel (iHex:EtOAc 3:1) aufgereinigt, um 125 mg der iodierten Verbindung **22** (0.21 mmol, 91 %) als farblosen Schaum zu liefern.

¹H-NMR (600 MHz, CDCl₃): δ/ppm = 8.23 (s, 1H, N-H), 8.02 (d, *J* = 2.4 Hz, 1H, 6-H), 5.99 (dd, *J* = 174.5, 6.6 Hz, 1H, 1'-H), 4.46 (dq, *J* = 143.2, 7.7 Hz, 1H, 3'-H), 4.14 (d, ¹*J*_{C-H} = 147.9 Hz, 1H, 5'-H), 3.92 – 3.63 (m, 2H, 4'-H, 5'-H), 2.50 (d, ¹*J*_{C-H} = 134.9 Hz, 1H, 2'-H), 2.27 (d, ¹*J*_{C-H} = 139.3 Hz, 1H, 2'-H), 1.15 – 1.00 (m, 28H, Si-CH, Si-CH(CH₃)₂).

¹³C-NMR (151 MHz, CDCl₃): δ/ppm = 159.79 (C-4), 149.52 (C-2), 144.24 (C-6), 85.45 (ddd, *J* = 43.0, 40.5, 2.1 Hz, C-4'), 84.87 (dt, *J* = 36.4, 2.2 Hz, C-1'), 67.21 (t, *J* = 39.2 Hz, C-3'), 60.55 (C-5), 59.96 (dd, *J* = 43.2, 2.6 Hz, C-5'), 40.10 (ddt, *J* = 38.3, 36.4, 2.2 Hz, C-2'), 17.92 (Si-CH(CH₃)₂), 17.80 (Si-CH(CH₃)₂), 17.62 (Si-CH(CH₃)₂), 17.43 (Si-CH(CH₃)₂), 17.26 (Si-CH(CH₃)₂), 17.16 (Si-CH(CH₃)₂), 17.13 (Si-CH(CH₃)₂), 17.01 (Si-CH(CH₃)₂), 13.66 (Si-CH(CH₃)₂), 13.17 (Si-CH(CH₃)₂), 12.95 (Si-CH(CH₃)₂), 12.62 (Si-CH(CH₃)₂).

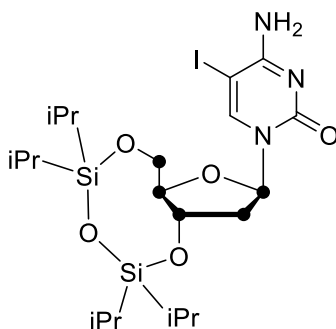
HRMS (ESI⁺): *m/z* berechnet für C₁₆¹³C₅H₃₈N₂O₆Si₂I⁺ [M+H]⁺ 602.1475; gefunden: 602.1481.

HRMS (ESI⁻): *m/z* berechnet für C₁₆¹³C₅H₃₆N₂O₆Si₂I⁻ [M-H]⁻ 600.1330; gefunden: 600.1334.

IR (ATR): ν (cm⁻¹) = 3184 (w), 3058 (w), 2945 (m), 2867 (m), 1681 (s), 1604 (w), 1463 (w), 1441 (w), 1385 (w), 1264 (m), 1100 (s), 1029 (s), 969 (m), 883 (s), 735 (m), 696 (s).

R_f (iHex:EtOAc 2:1): 0.57.

3',5'-(1,1,3,3,-Tetraisopropyldisiloxan-1,3-diyl)-5-iodo-2'-desoxy-(1',2',3',4',5'-¹³C₅)-cytidin (23)



In einem ausgeheizten Schlenk-Kolben wurden 110 mg des iodierten 2'-Desoxyuridin-Derivats **22** (0.19 mmol, 1.0 eq.) in 1.9 mL trockenem CH₂Cl₂ gelöst.

Anschließend wurden 270 μL NEt_3 (1.94 mmol, 10.0 eq.) und 5 mg DMAP (39 μmol , 0.2 eq.) zugesetzt und die Lösung wurde auf 0 °C gekühlt. Daraufhin wurden 88 mg 2,4,6-Triisopropylbenzolsulfonylchlorid (TPS-Cl, 0.29 mmol, 1.5 eq.) bei 0 °C in drei Portionen über einen Zeitraum von 30 min zugegeben, das Gemisch wurde auf Raumtemperatur erwärmen gelassen und für 15 h verrührt. Eine mit BuNH_2 behandelte Probe des Reaktionsgemisches zeigte dann dünnschichtchromatographisch vollständigen Umsatz des Startmaterials. Anschließend wurden 3.89 mL einer 0.5 M Lösung von NH_3 in 1,4-Dioxan (1.94 mmol, 10.0 eq.) zugetropft und es wurde 1 h bei Raumtemperatur verrührt. Dann wurde das Gemisch in 50 mL gesättigte wässrige NH_4Cl -Lösung gegossen und mit CH_2Cl_2 (3 x 50 mL) extrahiert. Die vereinigten organischen Phasen wurden über Na_2SO_4 getrocknet und *in vacuo* zur Trockene eingeengt. Aufreinigung des Rohprodukts mittels Säulenchromatographie an Silicagel ($\text{CH}_2\text{Cl}_2\text{:MeOH}$ 60:1 \rightarrow $\text{CH}_2\text{Cl}_2\text{:MeOH}$ 50:1 \rightarrow $\text{CH}_2\text{Cl}_2\text{:MeOH}$ 40:1) lieferte 60 mg des Aminierungsproduktes **23** (0.10 mmol, 53 %) als gelbliches Glas.

$^1\text{H-NMR}$ (600 MHz, CDCl_3): δ/ppm = 8.08 (d, J = 2.6 Hz, 1H, 6-H), 6.46 (br s, 1H, - NH_2), 5.99 (dd, J = 175.8, 6.7 Hz, 1H, 1'-H), 5.39 (br s, 1H, - NH_2), 4.37 (d, $^1J_{\text{C-H}}$ = 141.8 Hz, 1H, 3'-H), 4.31 – 3.63 (m, 3H, 4'-H, 5'-H), 2.52 (d, $^1J_{\text{C-H}}$ = 134.3 Hz, 1H, 2'-H), 2.34 (d, $^1J_{\text{C-H}}$ = 138.1 Hz, 1H, 2'-H), 1.13 – 0.91 (m, 28H, Si-CH, Si-CH(CH_3)₂).

$^{13}\text{C-NMR}$ (151 MHz, CDCl_3): δ/ppm = 163.71 (C-4), 154.75 (C-2), 146.80 (C-6), 85.73 – 84.91 (m, C-1', C-4'), 66.65 (t, J = 39.1 Hz, C-3'), 59.81 (d, J = 43.2 Hz, C-5'), 55.18 (C-5), 40.05 (ddt, J = 38.6, 36.6, 2.3 Hz, C-2'), 17.93 (Si-CH(CH_3)₂), 17.83 (Si-CH(CH_3)₂), 17.64 (Si-CH(CH_3)₂), 17.44 (Si-CH(CH_3)₂), 17.22 (Si-CH(CH_3)₂), 17.16 (2 x Si-CH(CH_3)₂), 17.03 (Si-CH(CH_3)₂), 13.65 (Si-CH(CH_3)₂), 13.17 (Si-CH(CH_3)₂), 12.91 (Si-CH(CH_3)₂), 12.59 (Si-CH(CH_3)₂).

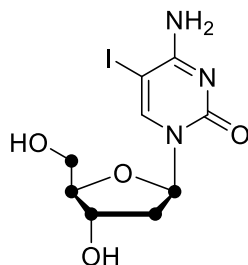
HRMS (ESI⁺): m/z berechnet für $\text{C}_{16}^{13}\text{C}_5\text{H}_{39}\text{N}_3\text{O}_5\text{Si}_2\text{I}^+$ $[\text{M}+\text{H}]^+$ 601.1635; gefunden: 601.1640.

HRMS (ESI⁻): m/z berechnet für $\text{C}_{16}^{13}\text{C}_5\text{H}_{37}\text{N}_3\text{O}_5\text{Si}_2\text{I}^-$ $[\text{M}-\text{H}]^-$ 599.1490; gefunden: 599.1495.

IR (ATR): ν (cm^{-1}) = 3083 (br w), 2944 (m), 2867 (m), 2360 (w), 2322 (w), 1631 (s), 1485 (s), 1465 (s), 1281 (w), 1246 (w), 1162 (w), 1120 (m), 1097 (s), 1047 (s), 1033 (s), 970 (w), 884 (m), 771 (w), 698 (m).

R_f (CH₂Cl₂:MeOH 40:1): 0.30.

5-Iodo-2'-desoxy-(1',2',3',4',5'-¹³C₅)-cytidin (24**)**



Zur Entschützung wurden 54 mg der Silyl-geschützten Verbindung **23** (90 μ mol, 1.0 eq.) in 3.50 mL trockenem CH₂Cl₂ und 1.40 mL trockenem EtOAc gelöst und mit 32 μ L einer 70%igen Lösung von HF in Pyridin (1.35 mmol, 15.0 eq.) versetzt. Das Gemisch wurde daraufhin 19 h bei Raumtemperatur verrührt. Zu der erhaltenen Suspension wurden dann 397 μ L Methoxytrimethylsilan (2.88 mmol, 32.0 eq.) zugetropft, das Gemisch wurde bei Raumtemperatur 1 h weiter Rühren gelassen und schließlich *in vacuo* zur Trockene eingengt. Der Rückstand wurde in 10 mL ddH₂O aufgenommen, mit CH₂Cl₂ (5 x 10 mL) gewaschen und die wässrige Phase wurde lyophilisiert, um 29.2 mg der entschützten Verbindung **24** (75 μ mol, 91 %) als farblosen Feststoff zu liefern.

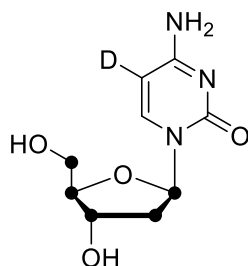
¹H-NMR (400 MHz, D₂O): δ /ppm = 8.30 (d, J = 2.7 Hz, 1H, 6-H), 6.16 (dt, J = 172.7, 6.3 Hz, 1H, 1'-H), 4.65 – 3.97 (m, 2H, 3'-H, 4'-H), 3.88 – 3.44 (m, 2H, 5'-H), 2.73 – 2.05 (m, 2H, 2'-H).

¹³C-NMR (101 MHz, D₂O): δ /ppm = 147.92 (C-6), 87.51 – 85.66 (m, C-1', C-4'), 69.73 (t, J = 36.4 Hz, C-3'), 60.57 (d, J = 41.5 Hz, C-5'), 39.48 (t, J = 36.1 Hz, C-2').

HRMS (ESI⁺): m/z berechnet für C₄¹³C₅H₁₃IN₃O₄⁺ [M+H]⁺ 359.0113; gefunden 359.0112.

HRMS (ESI⁻): m/z berechnet für C₄¹³C₅H₁₁IN₃O₄⁻ [M-H]⁻ 356.9968; gefunden: 356.9966.

5-Deutero-2'-desoxy-(1',2',3',4',5'-¹³C₅)-cytidin (**25**)



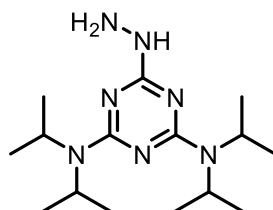
In einem ausgeheizten, evakuierten Schlenk-Kolben wurden 12.0 mg der entschützten Verbindung **24** (34 μ mol, 1.0 eq.), 23.0 mg NaHCO₃ (268 μ mol, 8.0 eq.) und 6.0 mg Pd/C mit 3.0 mL D₃COD versetzt und anschließend eine D₂-Atmosphäre erzeugt. Deuteriumgas wurde daraufhin für 5 min durch die Suspension blubbern gelassen und das Gemisch wurde 20 h bei Raumtemperatur verrührt. Anschließend wurde das Gemisch durch einen Spritzenvorsatzfilter gefiltert und *in vacuo* zur Trockene eingengt. Der Rückstand wurde in 4 mL ddH₂O aufgenommen, nochmals durch einen Spritzenvorsatzfilter filtriert und schließlich mittels *reversed phase HPLC* (0 % \rightarrow 5 % MeCN in H₂O in 45 min) aufgereinigt, um 5.1 mg des Deuterierungsproduktes **25** (22 μ mol, 65 %) als farblosen Feststoff zu erhalten.

¹H-NMR (400 MHz, D₂O): δ /ppm = 7.78 (d, J = 2.8 Hz, 1H, 6-H), 6.22 (dt, J = 171.0, 6.6 Hz, 1H, 1'-H), 4.37 (d, $^1J_{C-H}$ = 157.5 Hz, 1H, 3'-H), 4.22 – 3.48 (m, 3H, 4'-H, 5'-H), 2.36 (d, $^1J_{C-H}$ = 136.1 Hz, 1H, 2'-H), 2.26 (d, $^1J_{C-H}$ = 133.6 Hz, 1H, 2'-H).

HRMS (ESI⁺): m/z berechnet für C₄¹³C₅H₁₃DN₃O₄⁺ [M+H]⁺ 234.1209; gefunden: 234.1206.

HRMS (ESI⁻): m/z berechnet für C₄¹³C₅H₁₁DN₃O₄⁻ [M-H]⁻ 232.1064; gefunden: 232.1063.

6-Hydrazinyl-*N2,N2,N4,N4*-tetraisopropyl-1,3,5-triazin-2,4-diamin (**27**)



1.25 g Cyanurchlorid **26** (6.78 mmol, 1.0 eq.) wurden in 30 mL trockenem THF gelöst und auf 0 °C gekühlt. Eine Lösung aus 8.55 mL Diisopropylamin (61.01 mmol,

9.0 eq.) in 30 mL trockenem THF wurde langsam zugetropft, das Gemisch wurde 15 min bei 0 °C verrührt, anschließend auf Raumtemperatur und schließlich auf 78 °C erwärmt und daraufhin 7 Tage bei dieser Temperatur refluxiert. Die schmutzig-weiße Suspension wurde dann auf Raumtemperatur abkühlen gelassen, mit 200 mL H₂O versetzt und mit CH₂Cl₂ extrahiert (2 x 200 mL). Die vereinigten organischen Phasen wurden mit verdünnter wässriger Salzsäure (2 M, 200 mL) und anschließend mit gesättigter wässriger Natriumchloridlösung (200 mL) gewaschen, über Na₂SO₄ getrocknet und im *Anschluss in vacuo* zur Trockene eingedampft. Das Rohprodukt wurde mittels Säulenchromatographie an Silicagel (iHex:EtOAc 100:1) aufgereinigt, um 1.92 g eines weißen Feststoffs zu liefern, der direkt weiter eingesetzt wurde.

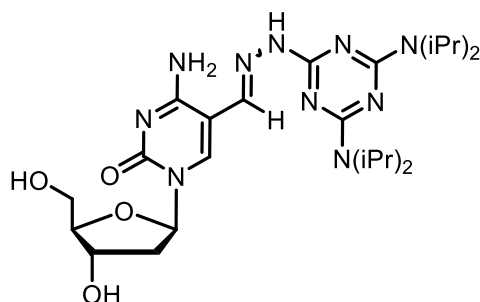
Das Zwischenprodukt (1.92 g) wurde in 25 mL THF gelöst und tropfenweise mit 3.09 mL Hydrazin-Monohydrat (50-60% in H₂O) versetzt. Das Gemisch wurde dann 24 h bei 78 °C refluxiert, anschließend auf Raumtemperatur abkühlen gelassen, mit 2.10 g frisch gepulvertem Kaliumhydroxid versetzt und 15 min bei Raumtemperatur verrührt. Daraufhin wurde mit CH₂Cl₂ extrahiert (2 x 200 mL), die vereinigten organischen Phasen über Na₂SO₄ getrocknet und Lösungsmittel *in vacuo* entfernt. Der Rückstand wurde via Säulenchromatographie an Silicagel (iHex:EtOAc 5:1 → iHex:EtOAc 2:1) aufgereinigt, um 2.07 g des Produktes (>quant.) al leicht gelbes Öl zu erhalten. Dieses wurde mit MeCN coevaporiert (3 x 30 mL) und anschließend lyophyllisiert, um das hydrazin-basierte Derivatisierungsreagenz **27** (1.61 g, 5.20 mmol, 77%) als leicht gelblichen Feststoff zu erhalten.

¹H-NMR (400 MHz, CDCl₃): δ/ppm = 6.11 (s, 1H, NH-NH₂), 3.84 (s, 2H, NH-NH₂), 1.43 – 1.04 (m, 28H, 3 x N(iPr)₂).

R_f (iHex:EtOAc 5:1): 0.35.

Weitere analytische Daten sind in der Literatur zu finden.^[118]

Derivatisierungsprodukt (28) von fdC mit 27



8.2 mg fdC (32.1 μmol , 1.0 eq.) wurden in 1000 μL ddH₂O gelöst und mit 20.0 mg Derivatisierungsreagenz **27** (64.3 μmol , 2.0 eq.) in 765 μL Methanol versetzt, wobei sich eine weiße Suspension bildete, die sich bei Zugabe von 176.5 μL Eisessig wieder auflöste. Das Reaktionsgemisch wurde 4 h bei 37 °C inkubiert und anschließend mit verdünnter wässriger Natronlauge (2 M) neutralisiert. Das Gemisch wurde lyophyllisiert, in 2.4 mL ddH₂O und 2.4 mL MeCN gelöst, durch einen Spritzenfilter gefiltert und mittels *reversed-phase* HPLC (50 % MeCN in ddH₂O isokratisch über 45 min) aufgereinigt, um das derivatisierte fdC **28** als weißen Feststoff (3.3 mg, 6.1 μmol , 19%, ca. 70:30 Gemisch aus (E)- und (Z)-Isomeren bei NMR) zu erhalten.

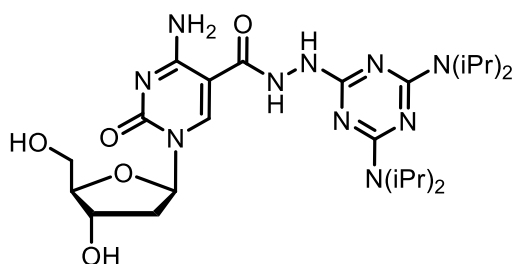
¹H-NMR (800 MHz, CD₃CN): δ/ppm = 9.04 (s, 0.7H, NH), 8.77 (s, 0.7H, NH), 8.52 (s, 0.3H, NH), 8.28 (s, 0.3H, CH=N-NH), 8.10 (s, 0.7H, CH=N-NH), 7.74 (2s, 1H, 6-H), 6.15 – 6.10 (m, 1H, 1'-H), 4.37 – 4.30 (m, 1H, 3'-H), 3.94 – 3.85 (m, 1H, 4'-H), 3.80 – 3.64 (m, 2H, 5'-H), 2.38 – 2.27 (m, 1H, 2'-H), 2.02 (dq, J = 4.4, 2.2 Hz, 1H, CHMe₂), 1.87 – 1.83 (m, 1H, CHMe₂), 1.42 – 1.19 (m, 24H CHMe₂).

¹³C-NMR (201 MHz, CD₃CN): δ/ppm = 163.94 (CH=N-NH), 145.03 (6-C), 88.37 (4'-C), 87.69 (1'-C), 71.01 (3'-C), 62.14 (5'-C), 41.66 (2'-C), 20.94 (CHMe₂).

HRMS (ESI⁺): m/z berechnet für C₂₅H₄₃N₁₀O₄⁺ [M+H]⁺ 547.3463; gefunden 547.3458.

HRMS (ESI⁻): m/z berechnet für C₂₅H₄₁N₁₀O₄⁻ [M-H]⁻ 545.3318; gefunden: 545.3320.

Derivatisierungsprodukt (29) von cadC mit 27



Das Triethylammoniumsalz von cadC (8.6 mg, 23.1 μmol , 1.0 eq.) wurde in 300 μL ddH₂O suspendiert, mit einer Lösung des Derivatisierungsreagenzes **27** (14.3 mg, 46.2 μmol , 2.0 eq.) in 600 μL MeCN, mit einer Lösung von HOBt (7.1 mg, 46.2 μmol , 2.0 eq.) in 300 μL ddH₂O und 300 μL MeCN und schließlich mit einer Lösung von EDC-HCl (22.1 mg, 115.5 μmol , 5.0 eq.) in 300 μL ddH₂O schrittweise versetzt. Anschließend wurde das Gemisch 3 h bei 37 °C inkubiert und daraufhin lyophilisiert. Der Rückstand wurde in 2.0 mL ddH₂O und 2.0 mL MeCN gelöst, durch einen Spritzenvorsatzfilter gefiltert und mittels *reversed-phase* HPLC (50 % MeCN in ddH₂O isokratisch über 45 min) aufgereinigt, um 8.3 mg des derivatisierten cadC **29** (14.8 μmol , 64%) als weißen Feststoff zu erhalten.

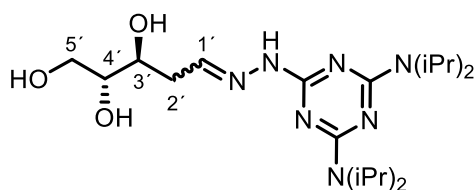
¹H-NMR (800 MHz, CD₃CN): δ/ppm = 8.76 (s, 1H, 6-H), 6.04 (t, J = 5.9 Hz, 1H, 1'-H), 4.30 (dt, J = 6.6, 4.9 Hz, 1H, 3'-H), 3.94 (td, J = 4.3, 3.1 Hz, 1H, 4'-H), 3.80 (dd, J = 12.4, 3.1 Hz, 1H, 5'-H), 3.67 (dd, J = 12.5, 4.1 Hz, 1H, 5'-H), 2.46 – 2.32 (m, 1H, 2'-H), 2.18 (dt, J = 13.8, 6.0 Hz, 1H, 2'-H), 1.39 – 1.01 (m, 28H, N(iPr)₂).

¹³C-NMR (151 MHz, CD₃CN): δ/ppm = 167.02 ((C=O)-NH-NH), 164.54, 164.22 (4-C), 155.24 (2-C), 144.40 (6-C), 87.76 (4'-C), 87.29 (1'-C), 69.63 (3'-C), 61.04 (5'-C), 41.19 (2'-C), 20.29 (N(iPr)₂).

HRMS (ESI⁺): m/z berechnet für C₂₅H₄₃N₁₀O₅⁺ [M+H]⁺ 563.3412; gefunden 563.3407.

HRMS (ESI⁻): m/z berechnet für C₂₅H₄₁N₁₀O₅⁻ [M-H]⁻ 561.3267 gefunden: 561.3268.

Derivatisierungsprodukt (**30**) von 2'-Desoxy-D-ribose mit **27**



5.1 mg 2'-Desoxy-D-ribose (38.0 μmol , 1.0 eq.) wurden in 1000 μL ddH₂O gelöst, mit 23.5 mg des Derivatisierungsreagenzes **27** (76.0 μmol , 2.0 eq.) in 765 μL Methanol versetzt und 176.5 μL Eisessig wurden hinzugegeben. Das Reaktionsgemisch wurde daraufhin 4 h bei 37 °C inkubiert, anschließend mit verdünnter wässriger Natronlauge (2 M) neutralisiert und schließlich lyophyllisiert. Der Rückstand wurde mittels Säulenchromatographie an Silicagel (CH₂Cl₂:MeOH 20:1 \rightarrow CH₂Cl₂:MeOH 8:1) aufgereinigt, um das Derivat **30** der 2'-Desoxy-D-ribose als weißen Feststoff (6.4 mg, 15.0 μmol , 40%) zu erhalten.

¹H-NMR (400 MHz, CD₃CN): δ/ppm = 7.36 (t, J = 4.7 Hz, 1H, 1'-H), 3.87 – 3.26 (m, 4H, 3'-H, 4'-H, 5'-H), 2.64 – 2.26 (m, 2H, 2'-H), 1.39 – 1.00 (m, 28H, N(iPr)₂).

¹³C-NMR (101 MHz, CD₃CN): δ/ppm = 165.25(C_{Ar}), 143.93 (CH=N-NH), 74.82 (4'-C), 71.67 (3'-C), 64.46 (5'-C), 36.10 (2'-C), 20.93 (N(iPr)₂).

HRMS (ESI⁺): m/z berechnet für C₂₀H₄₀N₇O₃⁺ [M+H]⁺ 426.3187; gefunden 423.3182.

R_f (CH₂Cl₂:MeOH 20:1): 0.10.

6.3. Untersuchungen der Aktivität von Sirtuinen auf die Deformylierung von fdC

6.3.1. Synthese eines DNA-Oligonukleotids mit isotopenmarkiertem fdC und *in vitro*-Experimente mit hSirt1

Insbesondere die Identifikation des oder der an der aktiven Deformylierung von genomischem fdC beteiligten Enzyme spielt für das tiefere Verständnis dieses Prozesses eine tiefgreifende Rolle. Ein Beispiel für eine Enzymklasse, die Deacetylierungen *in vivo* durchführen kann sind die sog. Histon-Deacetylasen (HDACs). Innerhalb der HDACs existieren 11 Vertreter, die in vier verschiedene Klassen eingeteilt werden können, wobei sich die Mitglieder der einzelnen Klassen in ihrer zellulären Lokalisation, ihrer mechanistischen Aktivität und natürlich in ihren spezifischen Funktionen unterscheiden. Nachdem *Seidel et al.* zeigen konnten, dass Sirtuine, die die Klasse 3 der HDACs bilden, neben ihrer typischen Deacetylase-Aktivität auch eine beachtliche Deformylase-Aktivität aufweisen können,^[121] rückten diese in den Fokus der Untersuchungen zur Deformylierung von fdC. Bei den Sirtuinen handelt es sich um NAD⁺-abhängige Enzyme, die unter Freisetzung einer Einheit Nikotinamid (NAM) Deacylierungs- bzw. Deformylierungsreaktionen katalysieren können. Ein hypothetischer Mechanismus einer Sirtuin-katalysierten Deformylierung von fdC ist in Abbildung 6-18 dargestellt.

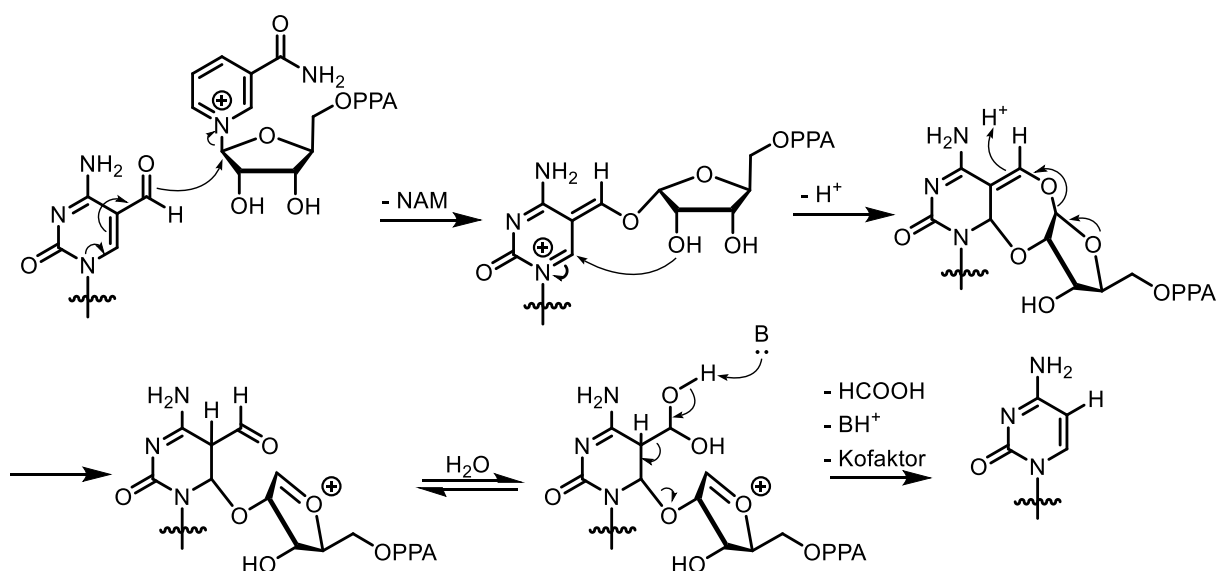


Abbildung 6-18: Postulierter Sirtuin-katalysierter Deformylierungsmechanismus von fdC.

Erste Untersuchungen zum Einfluss der Sirtuine auf die Deformylierung von fdC wurden von *Dr. Angie Kirchner* und *Dr. Katharina Iwan* durchgeführt. Dabei wurde die Deformylierungsaktivität von mESCs bei Verfütterungsexperimenten von F-fdC unter Zugabe des Sirtuin-Breitbandinhibitors NAM untersucht. Dabei wurde eine signifikant geringere Deformylierungsaktivität von F-fdC in den mit NAM behandelten Zellen im Vergleich zur unbehandelten Kontrolle ermittelt (Abbildung 6-19 A). Da es sich bei den Sirtuinen Sirt1 und Sirt6 um die einzigen Vertreter dieser Enzyme handelt, die im zellulären Nukleus lokalisiert sind, wurden zunächst mESC-KO-Zelllinien dieser Enzyme im Hinblick auf die Deformylierung von F-fdC untersucht. Die *in vivo* Experimente zeigten, dass bei einem Sirt1-KO eine signifikant reduzierte Deformylierungsaktivität von F-fdC in Vergleich zum Wildtyp beobachtet werden kann (siehe Abbildung 6-19 B).

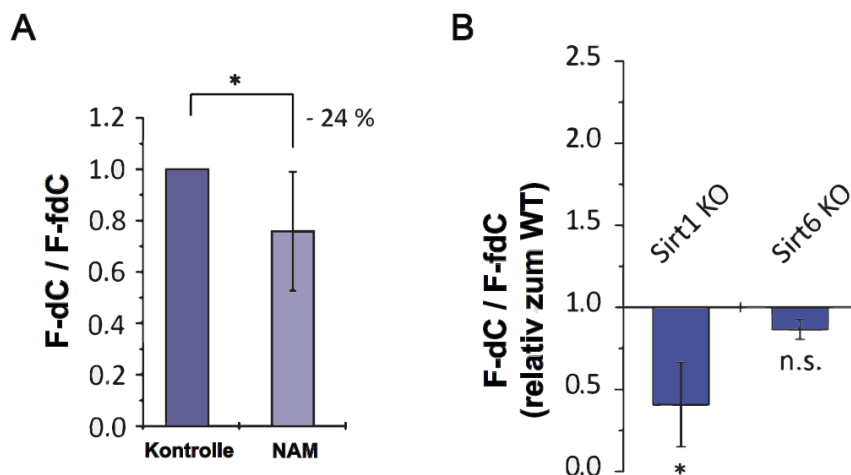


Abbildung 6-19: A) Beobachtete Deformylierungsaktivität von F-fdC in mESCs bei NAM-Behandlung im Vergleich zur normierten Kontrolle; B) Einfluss ausgewählter Sirtuin-KOs auf die *in vivo*-Deformylierung von F-fdC.

Zur genaueren Untersuchung der potentiellen Deformylase-Aktivität von Sirt1 auf fdC-enthaltende DNA sollten *in vitro* Studien mit Sirt1 durchgeführt werden. Dabei sollte ein DNA-Oligonukleotiddoppelstrang als Modellsystem eingesetzt werden, der isotope markiertes fdC enthalten sollte. Dazu wurde zunächst das isotope markierte $^{15}\text{N}_2$ -fdC Phosphoramidit zur Synthese der DNA-Oligonukleotide benötigt. Die Synthese beruht dabei auf einer Kombination von zwei bereits bekannten Synthesen. Zunächst wurde aus isotope markierten $^{15}\text{N}_2$ -Harnstoff **31** gemäß *Schiesser et al.* das isotope markierte, am Zucker geschützte fdC-Derivat **32** in einer sechsstufigen Synthese hergestellt.^[122] Anschließend wurde das Phosphoramidit **37** nach Vorschriften von *Schröder et al.*, die die Synthese des nicht-isotope markierten F-

fdC-Phosphoramidits beschreiben, durchgeführt.^[61] Da allerdings die $^{15}\text{N}_2$ -isotopenmarkierten Verbindungen der in Abbildung 6-20 gezeigten Syntheseroute nach dem geschützten fdC-Derivat **32** bisher nicht charakterisiert wurden, wird die Synthese hier kurz dargestellt und die Verbindungen im Experimentalteil dieses Kapitels (siehe Kapitel 6.3.3) näher charakterisiert.

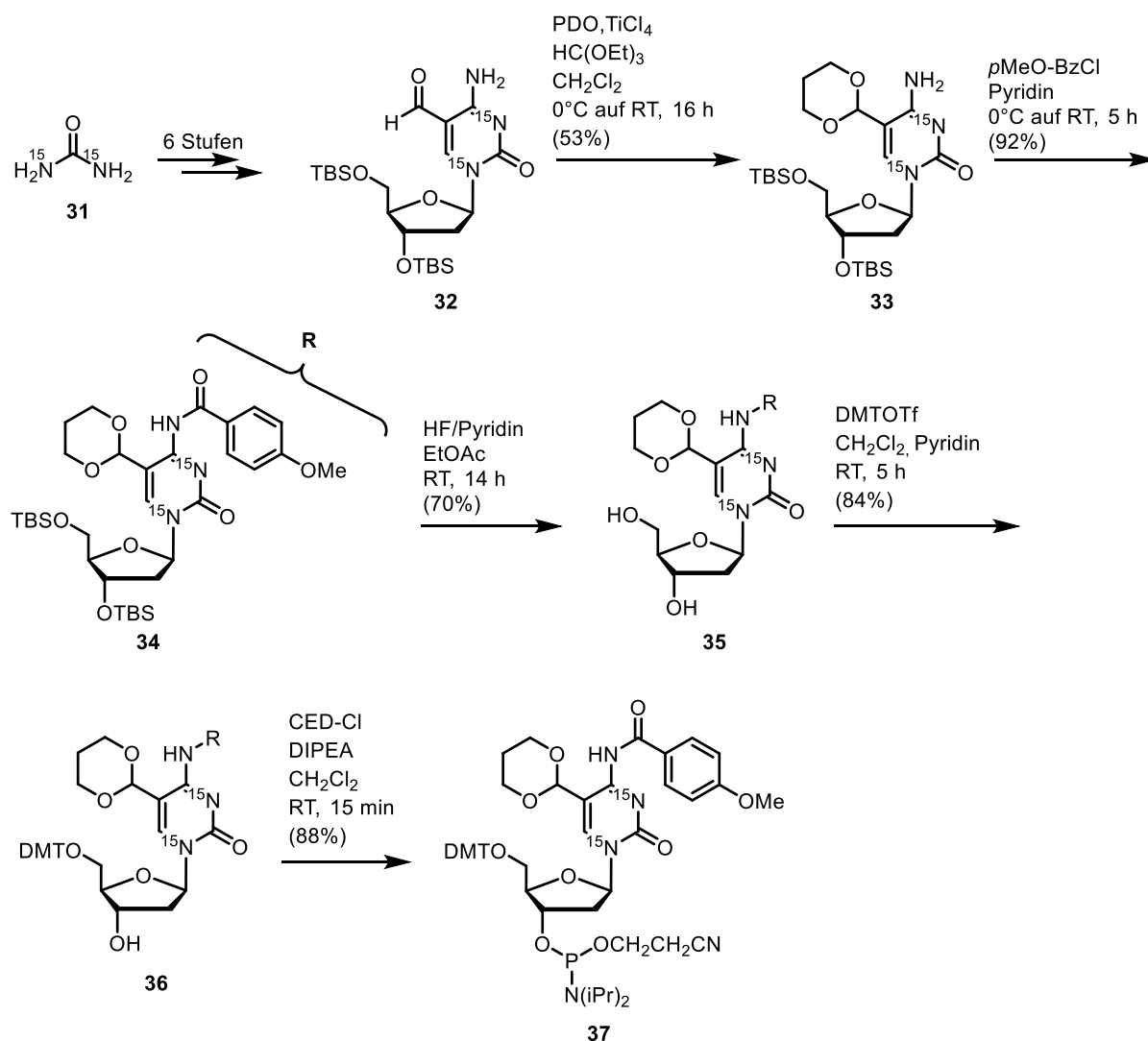


Abbildung 6-20: Syntheschema des $^{15}\text{N}_2$ -fdC-Phosphoramidits **37**.

Ausgehend vom geschützten fdC-Derivat **32** wurde zunächst die Aldehyd-Schutzgruppe durch Reaktion mit 1,3-Propandiol unter Lewis-sauren Bedingungen eingeführt. Dieses Dioxolan sollte sich nach der DNA-Festphasensynthese und der standardmäßigen Entschützung des Oligonukleotids unter basischen Bedingungen in einem zweiten Schritt unter Zusatz von Säure Oligonukleotid-kompatibel wieder abspalten lassen. Anschließend wurde dann durch Reaktion mit *para*-Methoxybenzoylchlorid die Schutzgruppe an der exozyklischen Aminofunktion des

Pyrimidinrings zu **34** eingeführt. Die elektronischen Eigenschaften dieser speziellen Schutzgruppe ermöglichen später die entsprechende Reaktivität des Phosphoramidites zur DNA-Festphasensynthese und gleichzeitig auch die nötige Labilität der Schutzgruppe gegenüber den Reaktionsbedingungen der abschließenden Entschützung. Daraufhin erfolgte die Entschützung des Nukleosids **34** von den Silyl-Schutzgruppen an den Hydroxygruppen der 3'- und 5'-Positionen mit Hydrogenfluorid in Pyridin, um anschließend selektiv an der 5'-Position die für die DNA-Festphasensynthese notwendige 4,4'-Dimethoxytrityl-Schutzgruppe (DMT-Schutzgruppe) unter Verwendung einer frisch hergestellten Lösung von 4,4'-Dimethoxytrityltriflat (DMT-OTf) in Dichlormethan zu installieren (**36**). Abschließend erfolgt die Phosphoramiditsynthese durch Reaktion mit Cyanoethyl-diisopropylchlorophosphoramidit (CED-Cl) und unter Zusatz von *Hünig*-Base in einer im Vergleich zur Literatur deutlich verkürzten Reaktionszeit zum isotopenmarkierten fdC-Phosphoramidit **37**. Bis auf die Ausbeute der Reaktion zur Einführung der Dioxolan-Schutzgruppe wurden im Verlauf dieser Syntheseroute bei den synthetischen Transformationen durchgehend annehmbare bis hohe Ausbeuten erreicht.

Nach erfolgreicher DNA-Festphasensynthese zum 30mer und Entschützung des Oligonukleotids unter den bekannten Bedingungen sollte dieses nun für *in vitro* Experimente mit Sirt1 verwendet werden. Die Experimente wurden von *Dr. Angie Kirchner* und *Ewelina Kaminska* und die massenspektrometrischen Messungen von *Dr. Katharina Iwan* durchgeführt, weswegen hier nur ein kurzer Überblick über das Experiment und das erhaltene Ergebnis gegeben werden soll. Ausführliche experimentelle Details sind in der Dissertation von *Dr. Angie Kirchner* zu finden.^[115]

Das fdC-enthaltende 30mer wurde zunächst zu Doppelsträngen hybridisiert, wobei das isotopenmarkierte fdC im CpG-Kontext entweder gegenüber einem dC oder einem mdC positioniert wurde. Anschließend wurden die Doppelstränge jeweils mit aktiven, humanem Sirt1 (hSirt1) für eine Stunde inkubiert, die DNA reisoliert, enzymatisch verdaut und schließlich das Verdaugemisch massenspektrometrisch analysiert. Die Aktivität des eingesetzten Sirt1 wurde vor der Inkubation mit den DNA-Doppelsträngen mittels eines von *Seidel et al.* entwickelten Fluoreszenzassays überprüft und bestätigt.^[121] Falls es durch das hSirt1 zu einer Deformylierung des isotopenmarkierten fdC am DNA-Doppelstrang käme, könnte bei der abschließenden

massenspektrometrischen Analytik das isoto­penmarkierte Deformylierungsprodukt $^{15}\text{N}_2\text{-dC}$ im Verdaugemisch detektiert werden (siehe Abbildung 6-21).

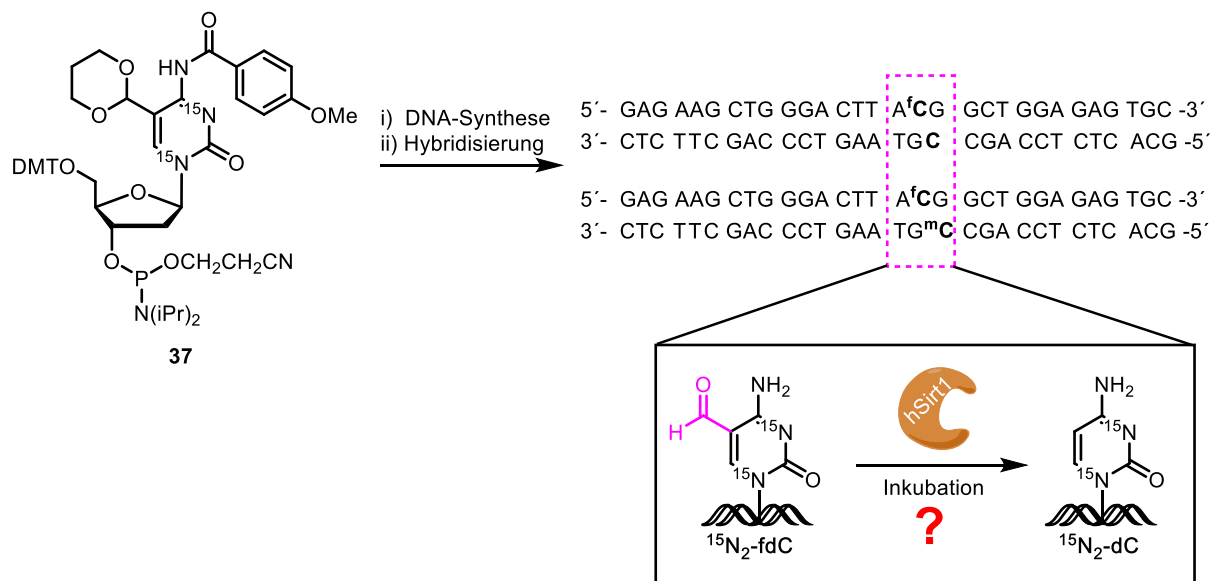


Abbildung 6-21: Überblick über die *in vitro*-Experimente unter Verwendung von Oligonukleotiden mit isoto­penmarkiertem fdC.

Allerdings konnte bei den massenspektrometrischen Untersuchungen sowohl beim Einsatz von 1 μM hSirt1 als auch bei der Inkubation der DNA-Doppelstränge mit 5 μM hSirt1 kein signifikantes Signal des isoto­penmarkierten $^{15}\text{N}_2\text{-dC}$ detektiert werden. Zwar wurde ein niedriges Signal der gesuchten Verbindung gemessen, allerdings war dieses in allen Proben zu finden, unabhängig, ob diese mit hSirt1 inkubiert wurden oder nicht. Auf Grund der geringen Signalintensität war das detektierte Signal wahrscheinlich dem natürlichen Isotopenhintergrund der untersuchten Verbindung zuzuordnen. Auch eine Verlängerung der Inkubationsdauer der DNA-Doppelstränge mit aktivem hSirt1 auf 24 h konnte keine veränderten Ergebnisse liefern.

Basierend auf diesen Daten scheint das hSirt1, dessen Aktivität vor der Inkubation mit den Oligonukleotiden fluoreszenzspektroskopisch bestätigt werden konnte, fdC nicht direkt deformylieren zu können. Das könnte einerseits daran liegen, dass sich das fdC in einem für das hSirt1 unpassenden DNA-Sequenzkontext befand. Die Sequenz der untersuchten Oligonukleotide wurde hier in Anlehnung an eine Dnmt3b Methylierungssequenz entwickelt. Möglicherweise würde eine andere Sequenz zu einem anderen Ergebnis kommen. Andererseits könnten auch weitere Proteine in Kombination mit hSirt1 zur Deformylierung von genomischem fdC notwendig sein.

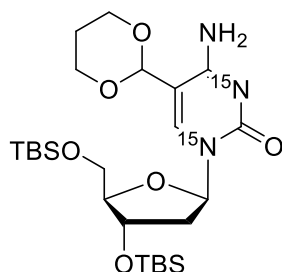
Dabei ist unter anderem bereits bekannt, dass Sirt1 in Kombination mit TETs oder Vertretern der Dnmt wirken kann.^[123] Allerdings sind hier zur Aufklärung des Einflusses von Sirt1 auf die aktive Deformylierung von genomischem fdC weitere Untersuchungen notwendig.

6.3.2. Projektbeiträge

Die *in vivo* Experimente zur Ermittlung des Einflusses von NAM auf die Deformylierung sowie die Studien der Sirtuin-KOs wurden von *Dr. Angie Kirchner* und *Dr. Katharina Iwan* durchgeführt. Experimentelle Details und ausführliche Diskussionen sind ihren Dissertationen zu entnehmen.^[115,116] Die Synthese des ¹⁵N₂-fdC-Phosphoramidits sowie die Entschützung des Oligonukleotids wurden von mir durchgeführt. Zur Synthese des Oligonukleotids mit dem von mir synthetisierten Phosphoramidit wurde die Firma *EllaBiotech* beauftragt. Die *in vitro* hSirt1-Studien wurden von *Ewelina Kaminska* und *Dr. Angie Kirchner*, die massenspektrometrischen Messungen von *Dr. Katharina Iwan* durchgeführt.

6.3.3. Experimentalteil

O-3',5'-Bis(tertbutyldimethylsilyl)-5-(1,3-dioxan-2-yl)-2'-desoxy-(M1,N3-¹⁵N₂)-cytidin (**33**)



Unter Ar-Atmosphäre wurden 321 mg des isotopenmarkierten fdC-Derivats^[122] **32** (0.66 mmol, 1.0 eq.) in 13.2 mL trockenem CH₂Cl₂ gelöst und auf 0 °C gekühlt. Anschließend wurden 132 µL Triethylorthoformiat (0.79 mmol, 1.2 eq.), 191 µL 1,3-Propandiol (2.64 mmol, 4.0 eq.) und schließlich 260 µL TiCl₄ in CH₂Cl₂ (1 M, 0.26 mmol, 0.4 eq.) bei 0 °C zugetropft. Das Gemisch wurde auf Raumtemperatur erwärmt und 16 h bei Raumtemperatur verrührt. Daraufhin wurde die erhaltene rotbraune Lösung in 150 mL eiskalte gesättigte, wässrige NaHCO₃-Lösung gegossen und mit EtOAc extrahiert (3 x 200 mL). Die vereinigten organischen Phasen wurden über Na₂SO₄ getrocknet, *in vacuo* eingeeengt und anschließend wurde das Rohprodukt durch Säulenchromatographie an Silicagel (CH₂Cl₂ → CH₂Cl₂:MeOH 200:1 → CH₂Cl₂:MeOH 100:1 → CH₂Cl₂:MeOH 80:1 → CH₂Cl₂:MeOH 60:1 → CH₂Cl₂:MeOH 50:1 → CH₂Cl₂:MeOH 30:1 → CH₂Cl₂:MeOH 20:1 → CH₂Cl₂:MeOH 15:1) um 189 mg des Acetal-geschützten Produkts **33** (0.35 mmol, 53 %) als farblosen Schaum zu erhalten.

¹H-NMR (400 MHz, CDCl₃): δ/ppm = 7.81 (s, 1H, 6-H), 6.25 (t, *J* = 6.1 Hz, 1H, 1'-H), 5.25 (s, 1H, CHOR₂), 4.33 (dt, *J* = 6.4, 3.4 Hz, 1H, 3'-H), 4.23 (m, 2H, O-CH₂-CH₂-CH₂-O), 3.96 – 3.65 (m, 5H, 4'-H, 5'-H, O-CH₂-CH₂-CH₂-O), 2.45 (dddd, *J* = 13.5, 6.2, 3.4, 1.5 Hz, 1H, 2'-H), 2.24 – 2.10 (m, 1H, O-CH₂-CH₂-CH₂-O), 1.96 (dtd, *J* = 13.3, 6.5, 2.6 Hz, 1H, 2'-H), 1.46 (ddd, *J* = 12.7, 2.8, 1.5 Hz, 1H, O-CH₂-CH₂-CH₂-O), 0.92 (s, 9H, C(CH₃)₃), 0.88 (s, 9H, C(CH₃)₃), 0.10 (s, 3H, SiCH₃), 0.10 (s, 3H, SiCH₃), 0.06 (s, 3H, SiCH₃), 0.05 (s, 3H, SiCH₃).

¹³C-NMR (101 MHz, CDCl₃): δ/ppm = 163.43 (d, *J* = 7.6 Hz, 4-C), 155.21 (2-C), 140.43 (d, *J* = 13.7 Hz, 6-C), 103.43 (5-C), 99.91 (CHOR₂), 87.95 (4'-C), 86.49 (1'-C), 72.03 (3'-C), 67.53 (O-CH₂-CH₂-CH₂-O), 67.50 (O-CH₂-CH₂-CH₂-O), 62.98 (5'-

C), 42.39 (2'-C), 26.18 (C(CH₃)₃), 25.91 (C(CH₃)₃), 25.69 (O-CH₂-CH₂-CH₂-O), 18.60 (C(CH₃)₃), 18.15 (C(CH₃)₃), -4.42 (SiCH₃), -4.74 (SiCH₃), -5.24 (SiCH₃), -5.30 (SiCH₃).

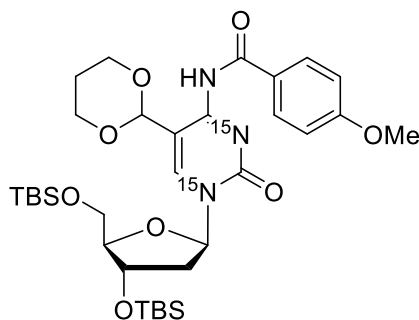
HRMS (ESI⁺): m/z berechnet für C₂₅H₄₈N¹⁵N₂O₆Si₂⁺ [M+H]⁺ 544.30147; gefunden: 544.3020.

HRMS (ESI⁻): m/z berechnet für C₂₅H₄₆N¹⁵N₂O₆Si₂⁻ [M-H]⁻ 542.2871; gefunden: 542.2880.

IR (ATR): ν (cm⁻¹) = 2952 (m), 2927 (m), 2855 (m), 1667 (s), 1506 (m), 1471 (s), 1253 (m), 1095 (s), 1030 (m), 1005 (m), 834 (s), 777 (s).

R_f (CH₂Cl₂:MeOH 20:1): 0.30.

O-3',5'-Bis(tertbutyldimethylsilyl)-N-4-(para-methoxybenzoyl)-5-(1,3-dioxan-2-yl)-2'-desoxy-(M1,N3-¹⁵N₂)-cytidin (34)



182 mg der Verbindung **33** (0.33 mmol, 1.0 eq.) wurden in 5.0 mL trockenem Pyridin gelöst und auf 0 °C gekühlt. Anschließend wurden 54 μ L para-Methoxybenzoylchlorid (0.40 mmol, 1.2 eq.) zugegeben, das Gemisch langsam auf Raumtemperatur erwärmt und 5 h bei dieser Temperatur gerührt. Nach vollständigem Umsatz des Startmaterials wurde das Reaktionsgemisch mit 100 mL gesättigter, wässriger NaHCO₃-Lösung versetzt und mit EtOAc extrahiert (3 x 100 mL). Die vereinigten organischen Phasen wurden über Na₂SO₄ getrocknet, *in vacuo* zur Trockene eingeeengt und der Rückstand wurde durch Säulenchromatographie an Silicagel (iHex:EtOAc 5:1 \rightarrow iHex:EtOAc 3:1) aufgereinigt, um 205 mg des an der aromatischen Aminogruppe einfach geschützten Produkts **34** (0.30 mmol, 92 %) als farbloses Wachs zu liefern.

¹H-NMR (600 MHz, CDCl₃): δ/ppm = 13.38 (d, *J* = 89.3 Hz, 1H, NH), 8.26 – 8.19 (m, 2H, (O=C)-C=CH), 7.97 (s, 1H, 6-H), 6.95 – 6.90 (m, 2H, CH=C-OCH₃), 6.25 (ddd, *J* = 7.4, 5.6, 1.3 Hz, 1H, 1'-H), 5.82 (s, 1H, CHOR₂), 4.41 (dt, *J* = 5.2, 2.3 Hz, 1H, 3'-H), 4.27 – 4.20 (m, 2H, O-CH₂-CH₂-CH₂-O), 4.03 (m, 2H, O-CH₂-CH₂-CH₂-O), 3.98 (dd, *J* = 3.7, 3.4 Hz, 1H, 4'-H), 3.87 (s, 3H, OCH₃), 3.81 – 3.76 (m, 2H, 5'-H), 2.42 – 2.30 (m, 1H, 2'-H), 2.27 – 2.11 (m, 1H, O-CH₂-CH₂-CH₂-O), 2.07 – 1.95 (m, 1H, 2'-H), 1.46 (m, 1H, 2'-H), 0.93 (s, 9H, C(CH₃)₃), 0.89 (s, 9H, C(CH₃)₃), 0.12 (s, 3H, SiCH₃), 0.12 (s, 3H, SiCH₃), 0.08 (s, 3H, SiCH₃), 0.08 (s, 3H, SiCH₃).

¹³C-NMR (151 MHz, CDCl₃): δ/ppm = 164.71 ((O=C)-C=CH), 162.43 (CH=C-OCH₃), 157.32 (4-C), 147.56 (C-2), 139.21 (6-C), 132.98 ((O=C)-C=CH), 131.77 ((O=C)-C=CH), 113.74 (CH=C-OCH₃), 113.50 (CH=C-OCH₃), 96.16 (CHOR₂), 88.36 (4'-C), 86.36 (1'-C), 72.58 (3'-C), 67.77 (O-CH₂-CH₂-CH₂-O), 67.75 (O-CH₂-CH₂-CH₂-O), 63.14 (5'-C), 55.56 (OCH₃), 26.18 (C(CH₃)₃), 25.91 (C(CH₃)₃), 25.73 (O-CH₂-CH₂-CH₂-O), 18.59 (C(CH₃)₃), 18.17 (C(CH₃)₃), -4.52 (SiCH₃), -4.69 (SiCH₃), -5.22 (SiCH₃), -5.48 (SiCH₃).

¹⁵N-NMR (41 MHz, CDCl₃): δ/ppm = -230.77, -235.05.

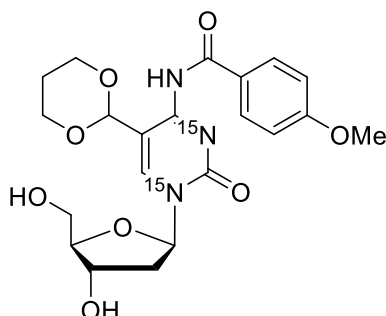
HRMS (ESI⁺): *m/z* berechnet für C₃₃H₅₄N¹⁵N₂O₈Si₂⁺ [M+H]⁺ 678.3385; gefunden: 678.3387.

HRMS (ESI⁻): *m/z* berechnet für C₃₃H₅₂N¹⁵N₂O₈Si₂⁻ [M-H]⁻ 676.3239; gefunden: 676.3246.

IR (ATR): ν (cm⁻¹) = 2953 (m), 2927 (m), 2855 (m), 1708 (s), 1651 (m), 1605 (m), 1566 (s), 1471 (w), 1330 (w), 1252 (s), 1161 (s), 1104 (s), 1030 (m), 1007 (w), 835 (s), 779 (m).

R_f (iHex:EtOAc 3:1): 0.47.

***N*-4-(Para-methoxybenzoyl)-5-(1,3-dioxan-2-yl)-2'-desoxy-(*N*1,*N*3-¹⁵N₂)-cytidin
(35)**



205 mg der Silyl-geschützten Verbindung **34** (0.30 mmol, 1.0 eq.) wurden in 4.5 mL trockenem EtOAc gelöst und mit 109 µL 70 % HF in Pyridin (4.54 mmol, 15.0 eq.) versetzt. Das Reaktionsgemisch wurde 14 h bei Raumtemperatur verrührt, die erhaltene Suspension mit 1.01 mL TMSOMe (9.68 mmol, 32.0 eq.) versetzt und eine weitere Stunde bei Raumtemperatur verrührt. Anschließend wurde der Niederschlag durch Zusatz von 20 mL MeOH gelöst und die organischen Lösungsmittel *in vacuo* entfernt. Der Rückstand wurde durch Säulenchromatographie an Silicagel (*dry load*, CH₂Cl₂:MeOH 30:1 → CH₂Cl₂:MeOH 20:1 → CH₂Cl₂:MeOH 15:1 → CH₂Cl₂:MeOH 10:1) aufgereinigt, um 94 mg des am Zucker entschützen Produkts **35** (0.21 mmol, 70 %) als farblosen Feststoff zu liefern.

¹H-NMR (800 MHz, DMSO-*d*₆): δ/ppm = 13.05 (d, *J* = 90.9 Hz, 1H, NH), 8.30 (s, 1H, 6-H), 8.19 – 8.09 (m, 1H (O=C)-C=CH), 7.95 – 7.78 (m, 1H, (O=C)-C=CH), 7.17 – 6.91 (m, 2H, CH=C-OCH₃), 6.15 (m, 1H, 1'-H), 5.74 (s, 1H, CHOR₂), 5.31 (d, *J* = 4.1 Hz, 1H, 3'-OH), 5.05 (s, 1H, 5'-OH), 4.27 4.21 (m, 1H, 3'-H), 4.17 – 3.91 (m, 4H, O-CH₂-CH₂-CH₂-O), 3.88 – 3.83 (m, 4H, 5'-H, OCH₃), 3.66 – 3.53 (m, 2H, 4'-H, 5'-H), 2.29 – 1.87 (m, 3H, 2'-H, O-CH₂-CH₂-CH₂-O), 1.52 – 1.40 (m, 1H, O-CH₂-CH₂-CH₂-O).

¹³C-NMR (201 MHz, DMSO-*d*₆): δ/ppm = 162.83 ((O=C)-C=CH), 156.38 (4-C), 140.41 (6-C), 131.63 ((O=C)-C=CH), 131.33 ((O=C)-C=CH), 129.64 ((O=C)-C=CH), 129.12 ((O=C)-C=CH), 113.80 (CH=C-OCH₃), 113.61 (CH=C-OCH₃), 95.03 (CHOR₂), 87.81 (4'-C), 85.52 (1'-C), 70.10 (3'-C), 66.85 (O-CH₂-CH₂-CH₂-O), 60.92 (5'-C), 55.44 (OCH₃), 40.25 (2'-C), 25.16 (O-CH₂-CH₂-CH₂-O).

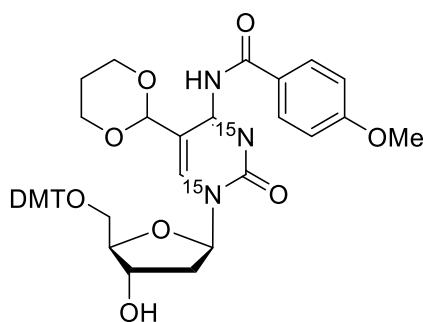
HRMS (ESI⁺): m/z berechnet für C₂₁H₂₆N¹⁵N₂O₈⁺ [M+H]⁺ 450.1655; gefunden: 450.1656.

HRMS (ESI⁻): m/z berechnet für C₂₁H₂₄N¹⁵N₂O₈⁻ [M-H]⁻ 448.1510; gefunden: 448.1513.

IR (ATR): ν (cm⁻¹) = 3335 (v), 1711 (m), 1662 (m), 1636 (m), 1606 (m), 1576 (s), 1515 (w), 1444 (m), 1397 (m), 1318 (w), 1252 (s), 1213 (s), 1184 (m), 1170 (w), 1098 (s), 1083 (s), 1067 (m), 1022 (m), 993 (s), 964 (w), 949 (s), 917 (m), 856 (m), 788 (s), 762 (s).

R_f (CH₂Cl₂:MeOH 10:1): 0.34.

O-5'-Dimethoxytrityl-N-4-(para-methoxybenzoyl)-5-(1,3-dioxan-2-yl)-2'-desoxy-(M1,N3-¹⁵N₂)-cytidin (36)



Im ausgeheizten Schlenkkolben wurden unter Ar-Atmosphäre 86 mg der Verbindung **35** (0.19 mmol, 1.0 eq.) in 1.90 mL trockenem CH₂Cl₂ und 1.90 mL trockenem Pyridin gelöst. Anschließend wurden tropfenweise 765 μ L einer frisch hergestellten 0.3 M Lösung von 4,4'-Dimethoxytrityltriflat in trockenem CH₂Cl₂ (0.23 mmol, 1.2 eq.) zugegeben und das Reaktionsgemisch wurde 5 h unter Lichtausschluss bei Raumtemperatur verrührt. Daraufhin wurde das Reaktionsgemisch in 50 mL gesättigte, wässrige NaHCO₃-Lösung gegossen und mit CH₂Cl₂ extrahiert (3 x 50 mL). Die vereinigten organischen Phasen wurden über Na₂SO₄ getrocknet, *in vacuo* eingeeengt, einmal mit Toluol (20 mL) coevaporiert und schließlich wurde der erhaltene Rückstand mittels Säulenchromatographie an Silicagel (CH₂Cl₂ → CH₂Cl₂:MeOH 60:1) aufgereinigt, um 120 mg des 5'-Dimethoxytrityl-geschützten Nukleosids **36** (0.16 mmol, 84 %) als gelblich-weißen Feststoff zu erhalten.

¹H-NMR (800 MHz, CD₂Cl₂): δ /ppm = 8.27 – 7.99 (m, 2H, (O=C)-C=CH), 7.49 – 7.46 (m, 2H, Ar_{DMT}-CH), 7.38 – 7.34 (m, 4H, Ar_{DMT}-CH), 7.33 – 7.30 (m, 2H, Ar_{DMT}-CH),

7.25 – 7.22 (m, 1H, Ar_{DMT}-CH), 6.97 – 6.94 (m, 2H, CH=C-OCH₃), 6.87 – 6.84 (m, 4H, Ar_{DMT}-CH), 6.27 (ddd, $J = 7.3, 5.8, 1.2$ Hz, 1H, 1'-H), 5.55 (s, 1H, CHOR₂), 5.33 (s, 1H, 6-H), 4.42 (dt, $J = 6.2, 3.1$ Hz, 1H, 3'-H), 4.08 – 4.02 (m, 1H, 4'-H), 3.92 – 3.88 (m, 1H, O-CH₂-CH₂-CH₂-O), 3.87 (s, 3H, OCH₃), 3.87 – 3.85 (m, 1H, O-CH₂-CH₂-CH₂-O), 3.78 (s, 3H, OCH₃), 3.78 (s, 3H, OCH₃), 3.57 – 3.52 (m, 2H, O-CH₂-CH₂-CH₂-O), 3.49 – 3.44 (m, 1H, 5'-H), 3.28 (dd, $J = 10.6, 3.9$ Hz, 1H, 5'-H), 2.48 – 2.42 (m, 1H, 2'-H), 2.23 (dddd, $J = 13.7, 7.5, 6.1, 2.5$ Hz, 1H, 2'-H), 1.43 (t, $J = 7.3$ Hz, 1H, O-CH₂-CH₂-CH₂-O).

¹³C-NMR (201 MHz, CD₂Cl₂): δ /ppm = 163.83 ((O=C)-C=CH), 159.24 (Ar-CH), 145.35 (Ar-CH), 136.39 (Ar-CH), 136.20 (Ar-CH), 130.67 (Ar-CH), 130.63 (Ar-CH), 128.65 (Ar-CH), 128.48 (Ar-CH), 127.37 (Ar-CH), 114.03 (CH=C-OCH₃), 113.72 (Ar-CH), 113.70 (Ar-CH), 87.25 (4'-C), 86.83 (1'-C), 72.61 (3'-C), 67.92 (O-CH₂-CH₂-CH₂-O), 67.89 (O-CH₂-CH₂-CH₂-O), 63.91 (5'-C), 56.02 (OCH₃), 55.78 (OCH₃), 41.82 (2'-C), 25.81 (O-CH₂-CH₂-CH₂-O).

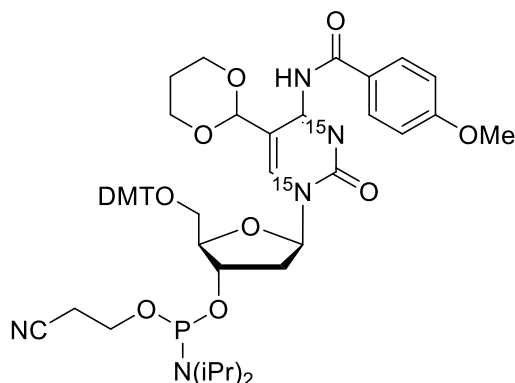
HRMS (ESI⁺): m/z berechnet für C₄₂H₄₄N¹⁵N₂O₁₀⁺ [M+H]⁺ 752.2962; gefunden: 752.2951.

HRMS (ESI⁻): m/z berechnet für C₄₂H₄₂N¹⁵N₂O₁₀⁻ [M-H]⁻ 750.2816; gefunden: 750.2824.

IR (ATR): ν (cm⁻¹) = 3309 (br w), 3052 (w), 2933 (w), 1705 (m), 1651 (m), 1605 (m), 1559 (s), 1507 (s), 1463 (m), 1442 (m), 1326 (w), 1245 (s), 1173 (s), 1160 (s), 1099 (s), 1030 (s), 1007 (w), 916 (w), 828 (m), 732 (s).

R_f (CH₂Cl₂:MeOH 20:1): 0.38.

O-3'-(Diisopropylcyanoethylphosphino)-O-5'-dimethoxytrityl-N-4-(para-methoxybenzoyl)-5-(1,3-dioxan-2-yl)-2'-desoxy-(M1,N3-¹⁵N₂)-cytidin (37**)**



Im ausgeheizten Schlenkkolben und unter Ar-Atmosphäre wurden 128 mg des Nukleosids **36** (0.17 mmol, 1.0 eq.) in 1.70 mL CH₂Cl₂ (HPLC-Güte) gelöst und bei Raumtemperatur mit 59 µL Diisopropylethylamin (0.34 mmol, 2.0 eq.) sowie 76 µL 2-Cyanoethyl-diisopropyl-chloro-phosphoramidit (CED-Cl, 0.34 mmol, 2.0 eq.) versetzt und 15 min verrührt. Daraufhin wurde das Reaktionsgemisch in 15 mL gesättigte, wässrige NaHCO₃-Lösung gegossen und mit CH₂Cl₂ extrahiert (1 x 20 mL). Die organische Phase wurde mit gesättigter, wässriger NaCl-Lösung gewaschen und über Na₂SO₄ getrocknet. Nach Einengen *in vacuo* wurde der Rückstand durch Säulenchromatographie an Silicagel (iHex:EtOAc 1:1 + 0.5 % Pyridin) aufgereinigt, um nach Coevaporation mit Toluol 141 mg des Phosphoramits **37** (0.15 mmol, 88 %) als farblosen Schaum zu liefern.

¹H-NMR (400 MHz, C₆D₆): δ/ppm = 8.71 (m, 2H, (O=C)-C=CH), 8.37 (s, 1H, 6-H), 7.71 - 7.68 (m, 2H, Ar_{DMT}-CH), 7.59 – 7.44 (m, 4H, Ar_{DMT}-CH), 7.30 – 7.25 (m, 2H, Ar_{DMT}-CH), 7.03 – 7.00 (m, 3H, Ar_{DMT}-CH), 6.87 (m, 6H, Ar_{DMT}-CH, CH=C-OCH₃), 6.39 (dt, *J* = 8.7, 4.4 Hz, 1H, 1'-H), 5.95 (s, 1H, CHOR₂), 4.53 (ddd, *J* = 9.3, 6.1, 2.8 Hz, 1H, 3'-H), 4.21 (dq, *J* = 24.7, 3.2 Hz, 1H, 4'-H), 3.75 – 3.61 (m, 3H, 5'-H, O-CH₂-CH₂-CH₂-O), 3.55 – 3.48 (m, 2H, O-CH₂-CH₂-CH₂-O), 3.45 - 3.38 (m, 2H, N-CH-(CH₃)₂), 3.37 (s, 3H, OCH₃), 3.36 (d, *J* = 2.2 Hz, 3H, OCH₃), 3.30 (ddd, *J* = 10.2, 6.2, 3.7 Hz, 1H, 5'-H), 3.25 (s, 3H, OCH₃), 3.23 – 3.09 (m, 2H, P-O-CH₂), 2.46 – 2.28 (m, 1H, 2'-H), 2.07 (dddd, *J* = 15.7, 10.6, 6.1, 3.1 Hz, 1H, 2'-H), 1.74 – 1.58 (m, 2H, CH₂-CN), 1.47 – 1.22 (m, 2H, O-CH₂-CH₂-CH₂-O), 1.11 – 1.02 (m, 9H, N-CH-(CH₃)₂), 0.98 – 0.89 (m, 3H, N-CH-(CH₃)₂).

¹³C-NMR (101 MHz, C₆D₆): δ/ppm = 163.79 (O=C)-C=CH, 159.41 (Ar-CH), 145.69 (Ar-CH), 141.01 (Ar-CH), 138.06 (6-C), 136.71 (Ar-CH), 136.35 (Ar-CH), 136.32 (Ar-CH), 132.89 (Ar-CH), 130.97 (Ar-CH), 129.49 (Ar-CH), 129.00 (Ar-CH), 128.73 (Ar-CH), 127.77 (Ar-CH), 125.86 (Ar-CH), 113.89 (Ar-CH), 113.87 (Ar-CH), 113.86 (Ar-CH), 113.84 (Ar-CH), 113.83 (Ar-CH), 96.43 (CHOR₂), 87.33 (4'-C), 86.12 (1'-C), 74.17 (3'-C), 67.56 (O-CH₂-CH₂-CH₂-O), 63.62 (5'-C), 58.89 (P-O-CH₂), 55.06 (OCH₃), 55.02 (OCH₃), 43.76 (N-CH-(CH₃)₂), 43.66 (N-CH-(CH₃)₂), 25.63 (O-CH₂-CH₂-O), 24.73 (N-CH-(CH₃)₂), 24.68 (N-CH-(CH₃)₂), 20.23 (CH₂-CN).

¹⁵N-NMR (41 MHz, C₆D₆): δ/ppm = -231.93, -234.68.

³¹P-NMR (162 MHz, C₆D₆): δ/ppm = 149.13, 148.09.

R_f (iHex: EtOAc 1:1 mit 1% NEt₃): 0.52.

Oligonukleotidsynthese

Für die Oligonukleotidsynthese mit dem synthetisierten ¹⁵N₂-fdC-Phosphoramidit **37** wurde die Firma *EllaBiotech* beauftragt, die auch die standartisierte basische Entschützung des Oligonukleotids nach der DNA-Festphasensynthese durchführte. Die saure Entschützung der Aldehyd-Funktion am Oligonukleotid wurde von mir unter Verwendung der bekannten Vorschriften von *Schröder et al.* durchgeführt.^[61] Allerdings wurden hier die Bedingungen zur Entschützung auf die Verwendung von 60% wässriger HOAc bei 20°C für 6 h abgemildert, um mögliche Nebenreaktionen am 30mer zu vermeiden. Trotz der milderer Entschützungsbedingungen konnte das Oligonukleotid in guter Ausbeute entschützt werden. Die erfolgreiche Entschützung wurde durch MALDI-TOF (gemessen von *Dr. Katharina Iwan*) bestätigt.

6.4. Entwicklung von Hydroxylamin-Reagenzien zur Sequenzierung von genomischem fdC mittels Technologien der 3. Generation

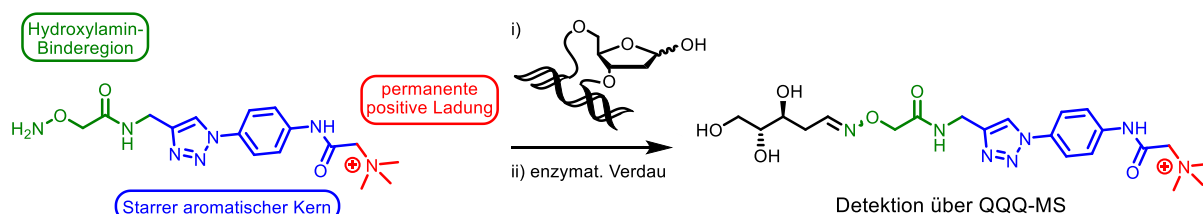
6.4.1. Design und Synthese der Verbindungen

Ein wesentlicher Aspekt, der zum tieferen Verständnis der fdC-Modifikationen im Genom beiträgt, ist die sequenzspezifische Detektion dieser modifizierten Basen über DNA-Sequenzierungsexperimente. Wie bereits weiter oben dargelegt, sind bereits Methoden zur Sequenzierung von genomischem fdC bekannt, diese beschränken sich allerdings auf die sog. Sequenzierungsmethoden der 2. Generation, die eine PCR-Amplifikation vor der eigentlichen Sequenzierung voraussetzen.^[69,70] Eine Weiterentwicklung dieser Sequenzierungstechniken hin zur Echtzeitsequenzierung auf Einzelmolekülniveau (Sequenzierung der 3. Generation) umgeht dieses mehrstufige Verfahren und könnte direkt Aufschluss über die fdC-Positionen im zu untersuchenden Genom liefern. Wie bereits aus Verfahren zur Echtzeitsequenzierung genomischen hmdCs bekannt ist, können durch Derivatisierung von hmdC mittels sterisch anspruchsvoller Reagenzien die hmdC-Positionen gezielt erkannt und in das Sequenzierungsergebnis integriert werden.^[77,78] Ein ähnliches Verfahren ist für die Sequenzierung von genomischem fdC denkbar, bisher aber nicht bekannt. Im Rahmen dieses Projekts sollte eine Bibliothek fdC-spezifischer, sterisch anspruchsvoller Derivatisierungsreagenzien entwickelt werden, die vor einer Echtzeitsequenzierung verwendet werden könnten, um die resultierenden fdC-Derivate selektiv und direkt detektieren zu können.

Ein Ansatzpunkt zum Design solcher Moleküle ist das von *Rahimoff et al.* vorgestellte Derivatisierungsreagenz **38** für abasische Stellen, welches mittels seiner Hydroxylaminfunktion selektiv mit Aldehyden reagieren kann.^[53] In der Literatur wurde gezeigt, wie die im Genom vorhandenen abasischen Stellen selektiv mit diesem Reagenz derivatisiert und anschließend über hochsensitive Massenspektrometrie detektiert werden konnten. Dabei weist das Hydroxylamin-Reagenz **38** eine starre, aromatische Kernstruktur auf und trägt eine permanente positive Ladung, die durch eine quartäre Ammoniumgruppe eingeführt wird (siehe Abbildung 6-22 A).^[53] Die aromatische Kernstruktur sowie die positive Ladung dienen dabei einerseits zu einer verbesserten Reaktivität der Verbindung **38** mit abasischen Stellen in der DNA, da die positive Ladung durch elektrostatische Wechselwirkung und das aromatische System durch nicht-kovalente π - π -Wechselwirkungen mit der

DNA interagieren können. Andererseits dient die im Reagenz implementierte positive Ladung insbesondere der verbesserten Detektion der derivatisierten abasischen Stellen, da die vor der Ionisation im Massenspektrometer bereits geladenen Verbindungen bei dieser Analytik zu besonders hoher Signalintensität führen.^[53] Weitere Experimente innerhalb des Arbeitskreises *Carell* zeigten, dass dieses Reagenz unter Zusatz von *p*-Anisidin als Katalysator auch mit fdC in DNA-Oligonukleotiden reagieren kann (Daten hier nicht gezeigt). Ausgehend von dem Design dieses Reagenzes sollte nun im Rahmen des Sequenzierungsprojekts eine Reihe von Verbindungen **39** ähnlicher Struktur abgeleitet und synthetisiert werden (Siehe Abbildung 6-22 B).

A



B

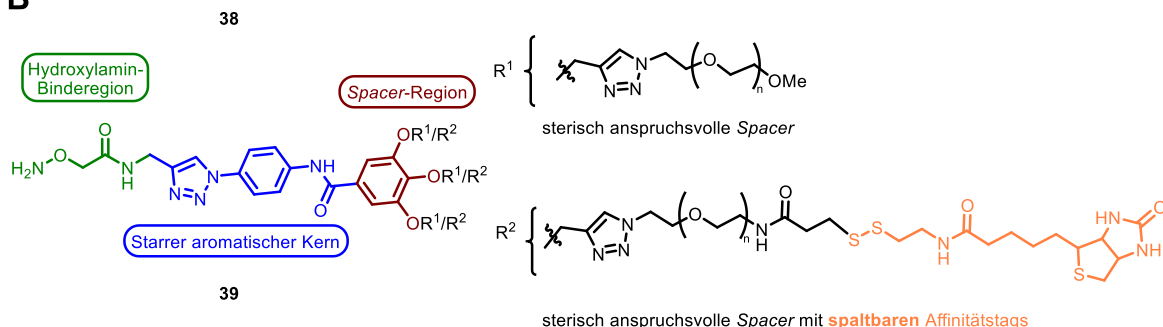


Abbildung 6-22: A) Darstellung des Hydroxylamin-basierten Reagenzes zur Derivatisierung abasischer Stellen; B) Darstellung der sterisch anspruchsvollen Hydroxylaminreagenzien zur möglichen Sequenzierung von genomischem fdC.

Dabei sollte die Hydroxylamin-Binderegion genauso erhalten bleiben, da bereits gezeigt werden konnte, dass diese in dem bisherigen Design nicht nur mit abasischen Stellen, sondern auch mit genomischem fdC reagieren kann. Ferner sollte auch die Kernregion zur π - π -Wechselwirkung mit den Nukleobasen der DNA beibehalten werden, an Stelle der quartären Ammoniumfunktion allerdings sollte nun die *Spacer-Region* eingeführt werden. Durch diese *Spacer-Region* sollten verschiedene sterisch anspruchsvolle Substituenten eingeführt werden, die schlussendlich einen Effekt auf die Echtzeitsequenzierung der derivatisierten fdC-Positionen haben sollten. Unter diesen sterisch anspruchsvollen Einheiten sollten

sowohl unfunktionalisierte *Spacer* auf Polyethylenglykol(PEG)-Basis (R^1) als auch mit selektiv abspaltbaren Affinitätstags funktionalisierte PEG-*Spacer* (R^2) untersucht werden. Als Affinitätstag für die *Spacer* der R^2 -Klasse wurde (+)-Biotin (im Folgenden nur „Biotin“ genannt) gewählt, welches spezifisch durch dessen Affinität zu Streptavidin angereichert werden kann. Dieses sollte über eine Disulfideinheit mit dem PEG-*Spacer* verbunden sein, um die Disulfidbrücke nach Anreicherung gezielt reduktiv spalten zu können. Generell wurden die PEG-basierten Moleküle für das Design der *Spacer* herangezogen, da die PEG-Einheiten die Wasserlöslichkeit des Gesamtsystems verbessern können, um eine akzeptable Wasserlöslichkeit der Zielmoleküle zu erreichen.

Die Synthese der Moleküle erfolgte mittels einer gemeinsamen Syntheseroute, bei der die Einheiten der *Spacer*-Region gegen Ende der Synthese durch eine Kupfer-katalysierte Alkin-Azid-Cycloadditionsreaktion (CuAAC-Reaktion) eingeführt wurden. Die gemeinsame Vorstufe **44** wurde in einer vierstufigen Synthese ausgehend von Hydroxylaminoessigsäure durchgeführt (siehe Abbildung 6-23).

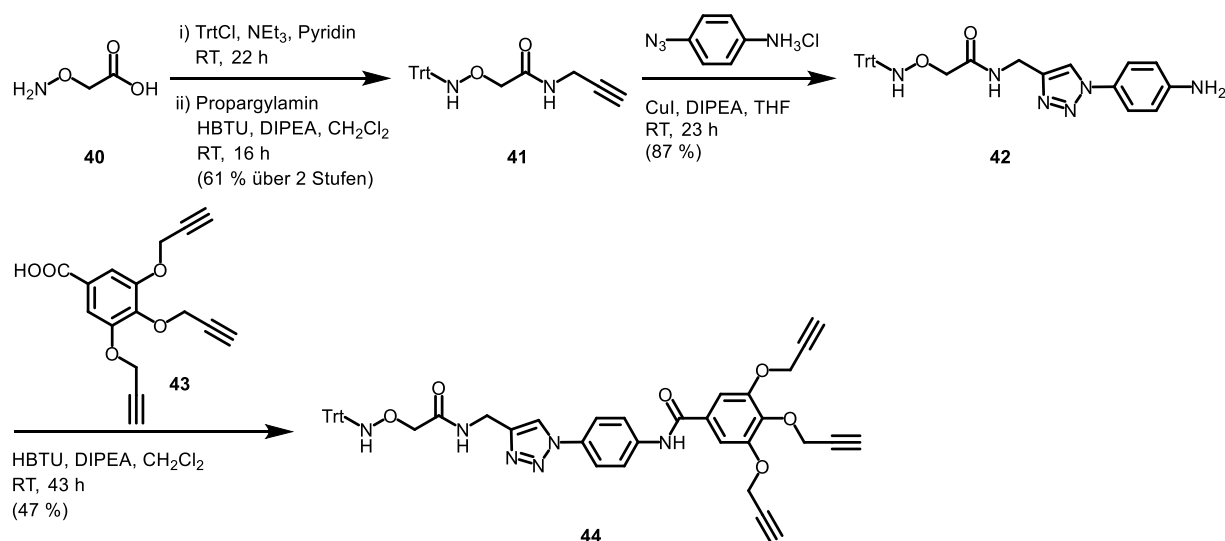


Abbildung 6-23: Schema zur Synthese des allgemeinen Vorläufermoleküls **44**.

Zunächst wurde die Hydroxylamin-Einheit der Hydroxylaminoessigsäure **40** selektiv Trityl-geschützt.^[124] Der Einsatz dieser säure-labilen Schutzgruppe ermöglicht die saure Entschützung der Hydroxylaminofunktion im letzten Schritt der Synthese unter Bedingungen, bei denen das Grundgerüst des Moleküls stabil ist.^[53] Nach Trityl-Schützung der Hydroxylaminfunktion wurde durch Amidkupplung mit Propargylamin unter Verwendung des Kupplungsreagenzes HBTU eine Alkineinheit eingeführt.^[53]

Diese Alkineinheit wurde im nächsten Schritt zur Kupfer-katalysierten Alkin-Azid-Cycloadditionsreaktion (CuAAC-Reaktion) mit 4-Azidoanilin-Hydrochlorid verwendet, um das aromatische Kerngerüst zu vervollständigen. Für diese CuAAC-Reaktion wurden verschiedene Reaktionsbedingungen untersucht, unter denen nur eine der untersuchten Bedingungen das gewünschte Produkt liefern konnte. Dabei handelte es sich um ein Vorgehen, das stöchiometrische Mengen an Kupfer(I)-iodid in sehr verdünnter Suspension in Tetrahydrofuran (0.01 M) in basischer Umgebung einsetzte.^[125] Scheinbar führten konzentriertere Reaktionsgemische oder die Verwendung anderer Kupferspezies zur Zersetzung des eingesetzten 4-Azidoanilins. Eine abschließende Amidkupplung des CuAAC-Reaktionsprodukts **42** mit einem dreifach alkinylierten Gallussäure-Derivat^[126] **43** unter erneuter Verwendung des Kupplungsreagenzes HBTU lieferte Produkt **44**, welches in der Folge an den Alkinpositionen funktionalisiert werden konnte.

Die unfunktionalisierten *Spacer* **47** konnten in einem einfachen zweistufigen Verfahren ausgehend von kommerziell erhältlichen PEG-Monomethylethern **45** synthetisiert werden. Hierzu wurde die freie Hydroxylgruppe zunächst tosyliert und in einem zweiten Schritt wurde die Abgangsgruppe dann durch eine S_N2-Reaktion mit Natriumazid substituiert.^[127] Die Synthese wurde beispielhaft für Diethylenglykol-Monomethylether und Triethylenglykol-Monomethylether in guten Ausbeuten durchgeführt.

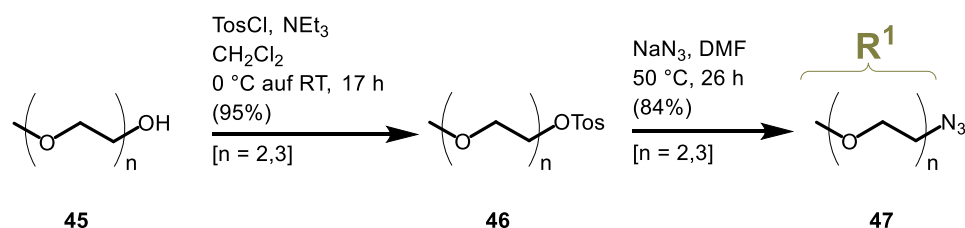


Abbildung 6-24: Schema der Synthese der unfunktionalisierten PEG-*Spacer* **47**.

Zur Synthese der biotinylierten *Spacer* wurde ein vierstufiges Verfahren ausgehend von Cystamin-Dihydrochlorid entwickelt. Zunächst wurde die Disulfidbrücke des Cystamins durch *m*CPBA zu einem asymmetrischen Sulfoxid oxidiert und anschließend wurde das Sulfoxid in einem zweiten Schritt durch β -Mercaptopropionsäure substituiert.^[128] Dadurch konnte aus dem symmetrischen Cystamin-Grundgerüst ein bifunktionales Molekül **49** synthetisiert werden, das sowohl eine terminale Carboxygruppe als auch eine terminale Aminofunktion aufwies. Dies wurde

im nächsten Schritt genutzt, um an der terminalen Aminogruppe gezielt eine Biotin-Einheit einzuführen. Dazu wurde der reaktive Ester des Biotins mit *N*-Hydroxysuccinimid (NHS) verwendet, welcher in einer einstufigen Synthese gemäß bekannter Literaturvorschriften hergestellt werden konnte.^[129] Der Einsatz des aktivierten Biotin-NHS-Esters verlangte dabei lediglich eine basische Reaktionsumgebung ohne den Zusatz von Kupplungsreagenzien, die unerwünschte Nebenreaktionen durch Aktivierung der Carboxylgruppe von **49** hervorrufen würden. Das erhaltene biotinylierte Disulfid **50** konnte im letzten Schritt durch eine Amidkupplung unter Verwendung von HATU als Kupplungsreagenz und dem kommerziell erhältlichen α -Azido-Triethylenglykol- ω -Amin in einen möglichen PEG-*Spacer* **51** der R²-Klasse überführt werden. Dieser sollte eine potentielle Abspaltung der über ein Disulfid verknüpften Biotineinheit ermöglichen. Alles in allem konnten die synthetischen Schritte bei der Herstellung des biotinylierten *Spacers* **51** in durchgehend annehmbaren Ausbeuten mit einer Gesamtausbeute von 30% ausgehend vom Cystamin-Dihydrochlorid durchgeführt werden.

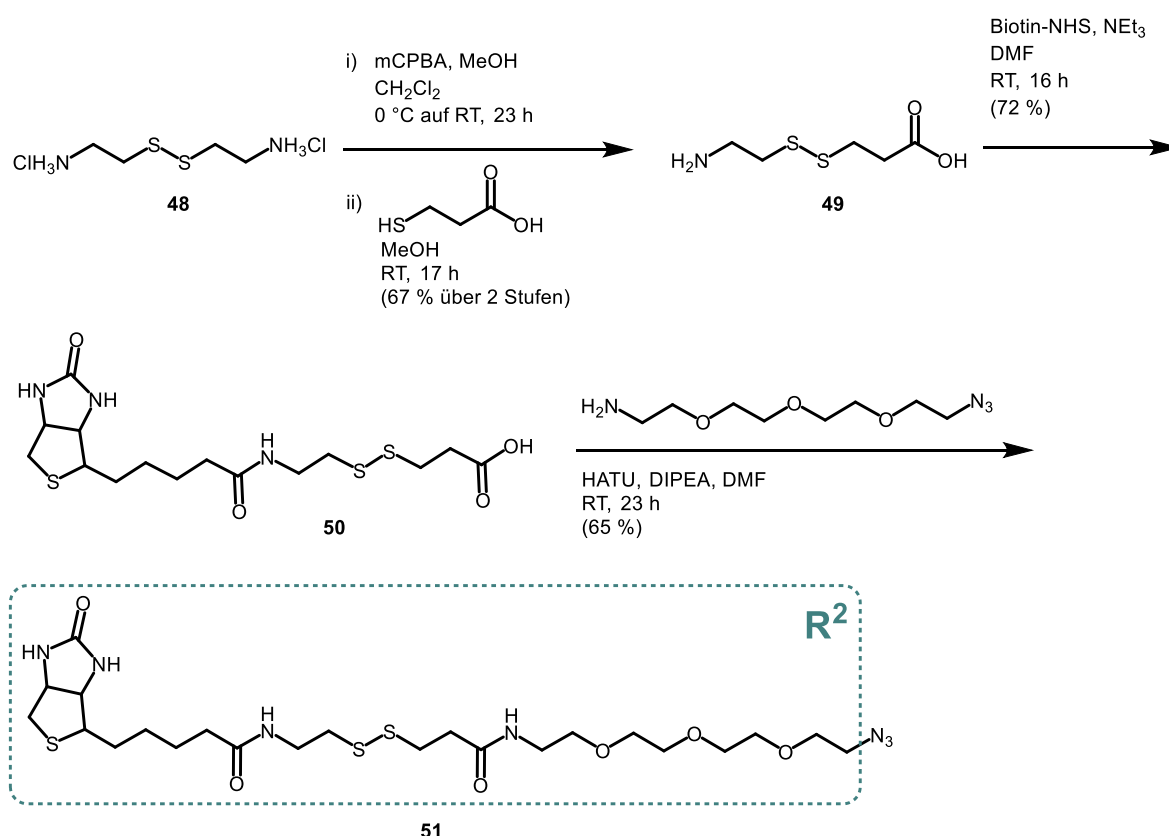


Abbildung 6-25: Schema der Synthese des Biotin-funktionalisierten *Spacers* **51**.

Ausgehend von der dreifach alkinylierten Verbindung **44** und den verschiedenen *Spacern* konnten nun die abschließenden Schritte zur Synthese der Zielmoleküle

durchgeführt werden. Dabei wurden sowohl Verbindungen **52** synthetisiert, die ausschließlich die unfunktionalisierten *Spacer* (R^1) aufwiesen, als auch Verbindungen **53**, die einen *Spacer* mit abspaltbarer Biotineinheit (R^2) trugen. Dabei wurden durch CuAAC-Reaktionen die *Spacer* unter Standardbedingungen an das dreifach alkylierte Grundgerüst addiert.^[53] Anschließend konnten die resultierenden Cycloadditionsprodukte ohne weitere Aufreinigung direkt entschützt werden. Dabei wurde die Entschützung von der Trityl-Gruppe an der terminalen Hydroxylamineinheit in einem Zweiphasengemisch aus wässriger Salzsäure (6 M) und Dichlormethan durchgeführt, wobei die wasserlöslichen Zielmoleküle von den organischen Verbindungen durch einfache Extraktion abgetrennt werden konnten.^[53] Eine Aufreinigung mittels semipräparativer *reversed-phase* HPLC lieferte die mit unfunktionalisierten PEG-Einheiten substituierten Hydroxylaminreagenzien **52** in guter Ausbeute (siehe Abbildung 6-26).

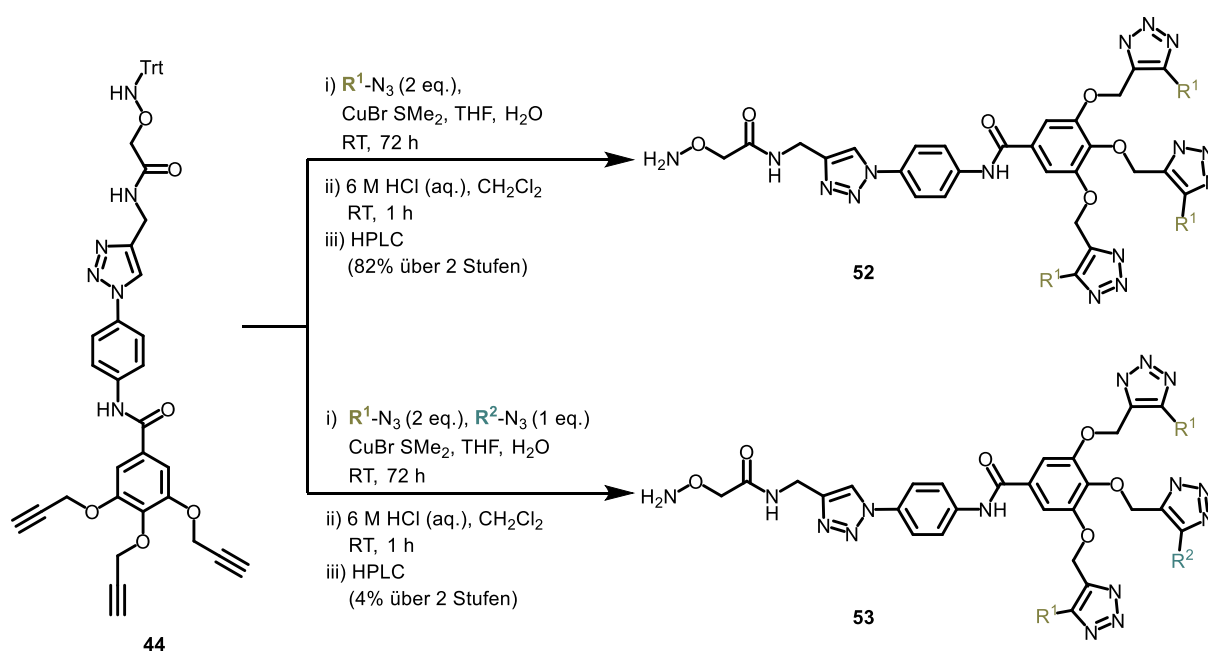


Abbildung 6-26: Schema der abschließenden Syntheseschritte zu den sterisch anspruchsvollen Hydroxylaminreagenzien.

Die einfach biotinylierte Verbindung **53** konnte allerdings nur in sehr schlechter Ausbeute isoliert werden. Dies kann einerseits am enormen sterischen Anspruch des biotinylierten *Spacers* **51** liegen, der nach CuAAC Reaktion mit **44** keinen vollständigen Umsatz zum dreifach Cycloadditions-Reaktionsprodukt ermöglichte. Andererseits reduzierte auch die Entstehung eines Produktgemisches durch den Einsatz zweier unterschiedlicher Azide in einem Eintopfverfahren für drei

Reaktivpositionen die zu erwartende Ausbeute. Nichtsdestotrotz konnte nach HPLC-Aufreinigung das symmetrische CuAAC-Reaktionsprodukt isoliert und für nachfolgende Experimente verwendet werden. Die Synthese und Isolation einer mehrfach biotinylierten Verbindung durch CuAAC von **44** mit ausschließlich biotinylierten *Spacern* führte zu keiner detektierbaren Produktbildung, was wiederum auf den sterischen Anspruch der biotinylierten *Spacer* zurückgeführt werden kann.

Neben den oben dargestellten PEG-substituierten Hydroxylaminreagenzien wurde auch eine sterisch weniger anspruchsvolle Verbindung synthetisiert. Dazu wurde zunächst 3,4,5-Trimethoxybenzoesäure **55** durch Esterhydrolyse von kommerziell erhältlichem Methyl-3,4,5-trimethoxybenzoat **54** in basischer Umgebung hergestellt, um die freie Carbonsäure anschließend unter Verwendung von HBTU mit 4-Azidoanilin zu koppeln. Daraufhin folgte eine CuAAC-Reaktion mit dem Tritylgeschützten Alkin **41**.^[53] Die geschützte Verbindung wurde anschließend durch das oben beschriebene Verfahren in einem Zweiphasengemisch aus Dichlormethan und wässriger Salzsäure (6 M) entschützt, und mittels *reversed-phase* HPLC aufgereinigt (siehe Abbildung 6-27).

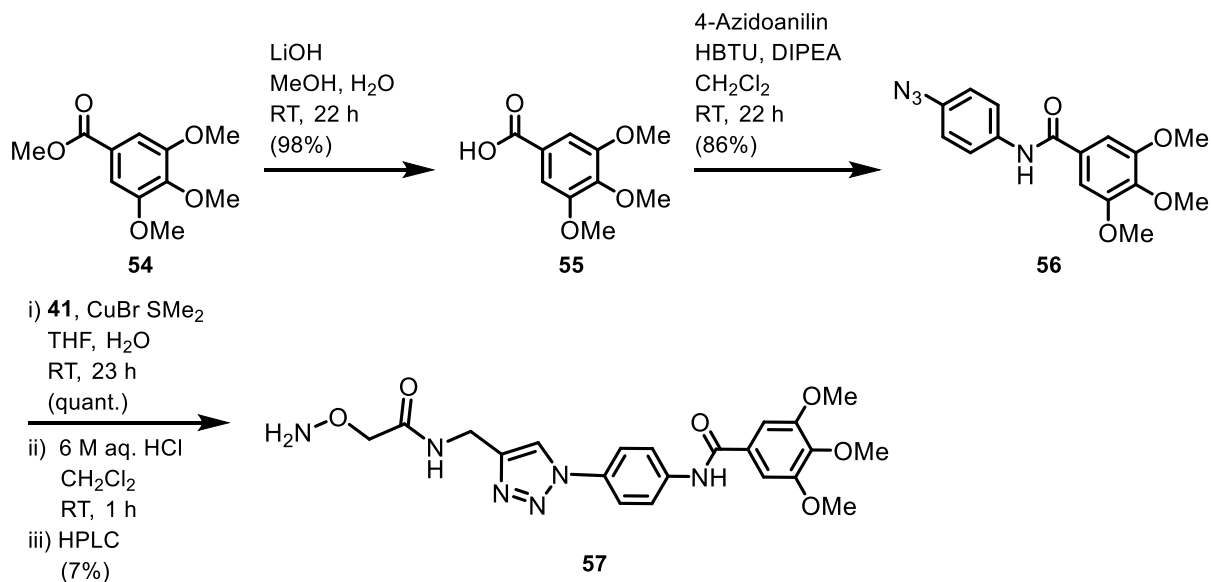


Abbildung 6-27: Syntheschema der Hydroxylaminverbindung **57**.

Verbindung **57**, die nur Methoxysubstituenten am Heterozyklus aufweist, sollte einen deutlich geringeren sterischen Anspruch als die zuvor beschriebenen Verbindungen **52** und **53** besitzen. Die geringe Ausbeute der Entschützungsreaktion ist hauptsächlich auf die generell schlechte Löslichkeit der geschützten Verbindung zurückzuführen. Bei der Reaktion im oben beschriebenen Zweiphasengemisch aus

Dichlormethan und wässriger Salzsäure kam es im Gegensatz zu den zuvor beschriebenen Prozessen (Abbildung 6-26) zu keiner klaren Phasentrennung und es wurde ebenfalls kein vollständiger Umsatz des Startmaterials detektiert. Dennoch konnte eine HPLC-Aufreinigung des teilweise umgesetzten Reaktionsgemisches das Hydroxylaminreagenz **57** in schlechter Ausbeute liefern.

Neben den bereits dargestellten Verbindungen wurde auch die Synthese einer regioselektiv funktionalisierbaren, strukturähnlichen Verbindung durchgeführt. Im Rahmen dieser Synthese sollte die abschließende Funktionalisierung des Grundgerüsts mit verschiedenen *Spacern* im Eintopfverfahren (siehe Abbildung 6-26) umgangen werden. Ein bifunktionelles Grundgerüst sollte dabei die Funktionalisierung mit zwei verschiedenen *Spacern* in zwei voneinander getrennten Syntheseschritten ermöglichen. Dieser aus organisch-chemischer Sicht elegantere Ansatz könnte insgesamt zu einer erhöhten Ausbeute und gleichzeitig zu einer einfacheren Aufreinigung der Endprodukte führen, da es zu keiner Bildung eines Produktgemisches wie bei dem Vorgehen im Eintopfverfahren kommen kann.

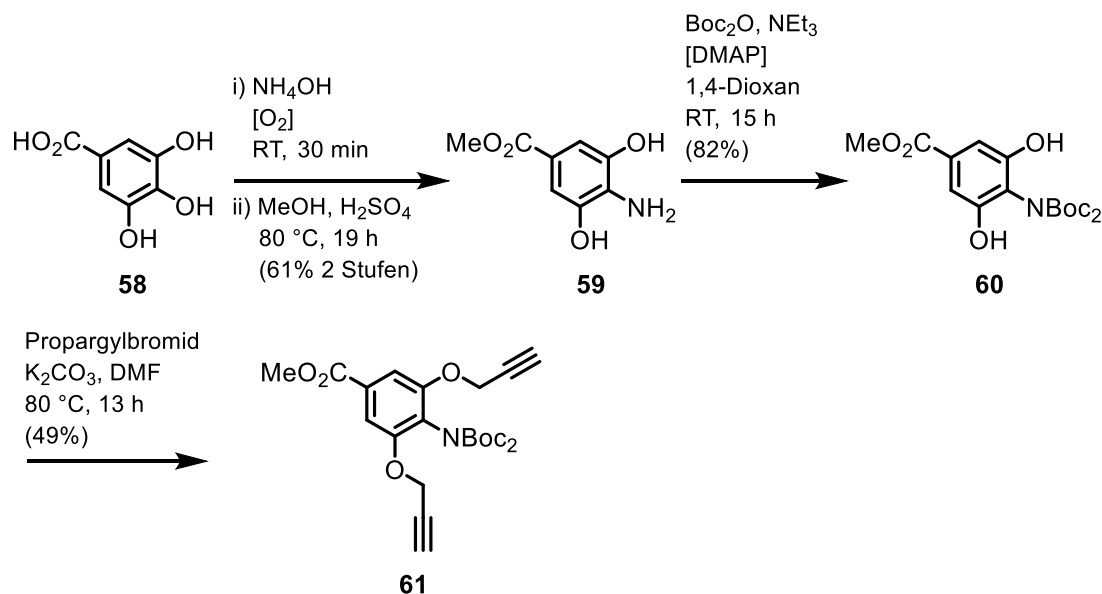


Abbildung 6-28: Syntheseroute zur multifunktionalisierbaren Verbindung **61**.

Zunächst wurde dazu Gallussäure **58** in der 4-Position mit konzentriertem wässrigem Ammoniak aminiert. Diese Aminierung wurde bereits im analytischen Maßstab von *Hashida et al.*^[130] beschrieben und hier im präparativ-synthetischen Maßstab durchgeführt. Dabei ist besonders zu beachten, dass molekularer Luftsauerstoff für eine erfolgreiche Reaktion ohne Schutzgasatmosphäre erforderlich ist, da es sich

grundsätzlich hier um einen redoxbasierten Vorgang handelt. Gemäß *Hashida et al.* kommt es hierbei zunächst zu einer Oxidation des pyrogallolischen Polyphenols Gallussäure **58** durch Luftsauerstoff zum *ortho*-Chinon, welches dann zum *ortho*-Chinonimin aminiert und schließlich durch andere Gallussäuremoleküle zum gewünschten Aminierungsprodukt reduziert wird.^[130] Ausgehend von der aminierten Verbindung wurde direkt ohne zwischenzeitliche Aufreinigung die Carboxygruppe unter *Fischer*-Bedingungen in Methanol zu **59** verestert, welches über Säulenchromatographie aufgereinigt werden konnte. Anschließend wurde die Aminogruppe der Verbindung selektiv mit Di-*tert*-butyl-dicarbonat mit *tert*-Butoxycarbonyl-Schutzgruppen (Boc-Schutzgruppen) zu **60** geschützt, um im weiteren Verlauf selektiv die Alkineinheiten einführen zu können. Dies wurde durch einfache S_N2-Reaktion von **60** mit Propargylamid in basischer Umgebung erreicht. Die somit erhaltene, symmetrische Basisverbindung **61** wies zwei alkinylierte Substituenten und eine geschützte Aminofunktion am aromatischen Ring auf (siehe Abbildung 6-28).

Im nächsten Schritt konnten die unfunktionalisierten PEG-*Spacer* **47** an die Alkineinheiten durch CuAAC-Reaktion addiert werden, um Verbindung **62** zu erhalten. Die Boc-geschützte Aminogruppe der Verbindung **62** sollte erst relativ am Ende der Synthesesequenz in einer separaten Reaktion mit einem biotinylierten *Spacer* funktionalisiert werden.

Ausgehend vom Cycloadditionsprodukt **62** wurde zunächst der Methylester der Verbindung zur Carbonsäure hydrolysiert, um anschließend die freigelegte Carboxygruppe unter Verwendung von HBTU mit 4-Azidoanilin zu kuppeln. Das erhaltene Kupplungsprodukt **64** wurde an der Aminofunktion von den Boc-Schutzgruppen durch ein Gemisch aus TFA in CH₂Cl₂ in annehmbarer Ausbeute zu **65** entschützt und in einem abschließenden Schritt mit dem Trityl-geschützten Alkin **41** in einer CuAAC-Reaktion umgesetzt,^[53] um das Cycloadditionsprodukt **66** zu erhalten (siehe Abbildung 6-29). Dieses Reaktionsprodukt weist eine freie Aminogruppe am aromatischen Kerngerüst auf, die selektiv unter Amidkupplungsbedingungen zur Reaktion gebracht werden kann. PEG-basierte *Spacer*, die eine terminale Carboxygruppe und einen abspaltbaren Affinitätstag enthalten, sind kommerziell erhältlich und wurden im Rahmen dieses Projekts nicht synthetisiert.

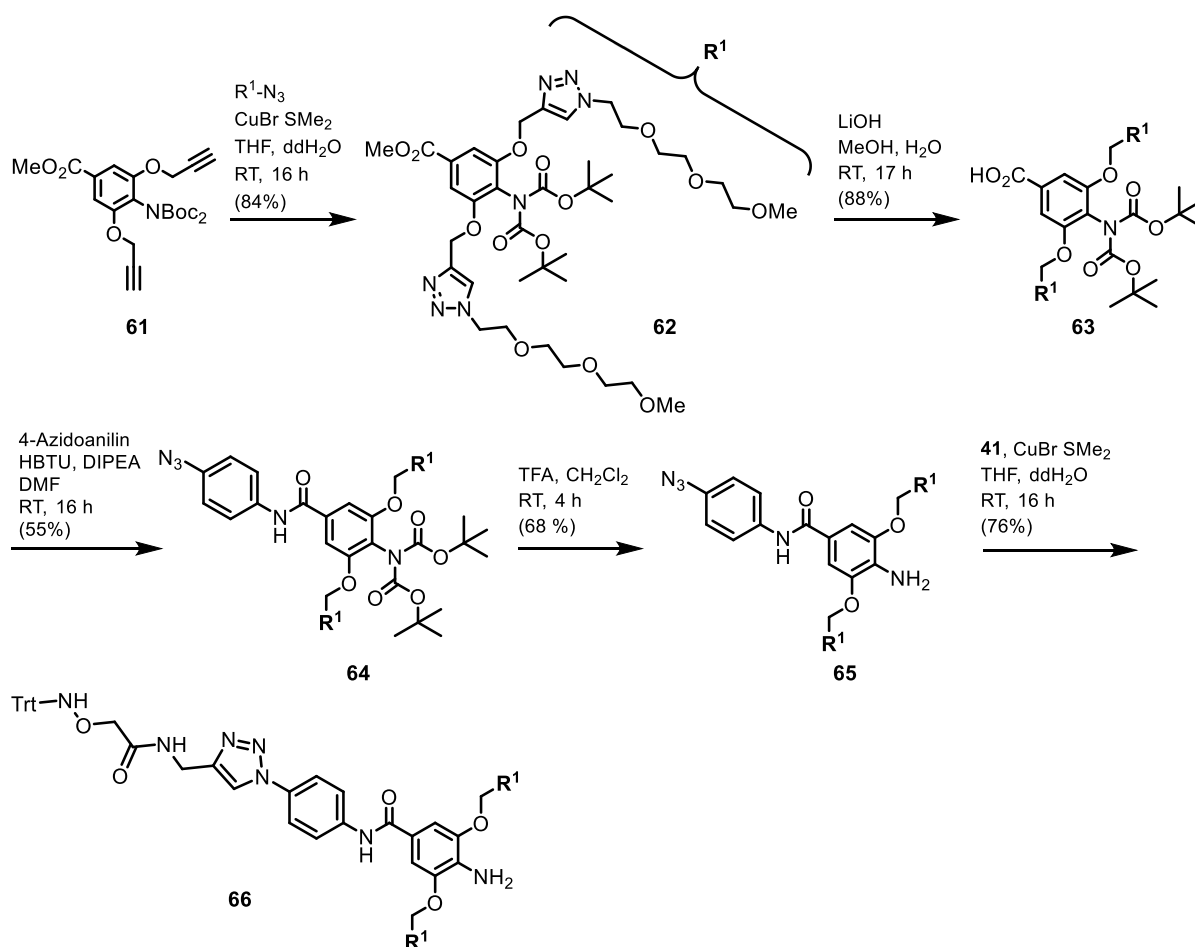


Abbildung 6-29: Schema zur Synthese des Vorläufers von selektiv bifunktionalisierbaren sterisch anspruchsvollen Hydroxylaminreagenzien.

Auf Grund der aufwendigen Synthese der Verbindung **66** wurde diese hier nicht weiter umgesetzt, sondern sollte erst nach ersten Ergebnissen von Sequenzierexperimenten weiter verwendet werden. Die Reaktion sollte dann mit einem geeigneten Affinitätstag-funktionalisierten *Spacer* erfolgen, der einen aus den ersten Experimenten ermittelten passenden sterischen Anspruch aufweist. Die folgende Entschützung der Tritylschutzgruppe sollte unter den oben beschriebenen Bedingungen möglich sein und könnte das regiochemisch selektiv funktionalisierte Hydroxylaminreagenz möglicherweise in besserer Ausbeute und einfacherer Aufreinigung als die oben beschriebene Syntheseroute (siehe Abbildung 6-26) liefern.

Ausgehend von der hier synthetisierten Gruppe an sterisch anspruchsvollen Hydroxylaminreagenzien, sollte im nächsten Schritt die Reaktion der Verbindungen mit fdC auf Oligonukleotidebene untersucht werden, da die Verwendung der Verbindungen auf deren Reaktion mit genomischem fdC ausgerichtet ist.

6.4.2. Reaktion der Derivatisierungsreagenzien mit fdC-enthaltenden Oligonukleotiden

Zur Untersuchung der Reaktion der synthetisierten Hydroxylaminreagenzien mit fdC auf Oligonukleotidebene, wurde ein synthetisches DNA-11mer mit einer fdC-Modifikation im CpG-Kontext verwendet. Das Design der durchgeführten Experimente beruht auf Studien von *Dr. Matthias Kurz*, der die Reaktion des bekannten Hydroxylaminreagenzes **38** mit fdC-enthaltenden DNA-Oligonukleotiden untersuchte. Dabei wurde entdeckt, dass zur Produktbildung Anisidin als nukleophiler Katalysator zugesetzt werden musste. Dieses musste in den hier durchgeführten Studien ebenfalls eingesetzt werden, da die hier verwendeten, sterisch anspruchsvollen Derivatisierungsreagenzien ein noch langsames Reaktionsverhalten mit fdC als das bekannte Hydroxylaminreagenz **38** aufwiesen (Daten hier nicht gezeigt). Zur gezielten Untersuchung der Reaktion wurde das Oligonukleotid in gepufferter wässriger Lösung mit Anisidin und einem der synthetisierten Hydroxylaminreagenzien bei 37 °C inkubiert und Proben nach definierten Zeitpunkten entnommen. Die Reaktion wurde dabei durch Zugabe eines Überschusses an Aceton gestoppt. Die Proben wurden dann jeweils mittels *reversed-phase* HPLC analysiert, um den Umsatz der eingesetzten DNA zu untersuchen. Abbildung 6-30 zeigt beispielhafte Chromatogramme ($\lambda = 260$ nm) der Reaktion des fdC-enthaltenden Oligonukleotids mit einem Überschuss an Hydroxylaminreagenz **67** in Gegenwart von Anisidin beim Reaktionsstart (5 min Inkubation), nach 6 h und nach 24 h Inkubation unter den optimierten Reaktionsbedingungen.

Anhand der Chromatogramme ist erkennbar, dass nach einer Inkubationszeit von 6 h bereits Produktbildung deutlich erkennbar und nach 24 h vollständiger Umsatz des Startmaterials zum Oximprodukt detektiert werden konnte. Eine massenspektrometrische Untersuchung des Reaktionsproduktes mittels MALDI-TOF konnte die Bildung des gewünschten Oximproduktes bestätigen. Anhand der Chromatogramme ist auch zu erkennen, dass die Reaktion unter diesen optimierten Bedingungen sauber und selektiv ohne Nebenproduktbildung ablief. Neben dem untersuchten Hydroxylaminreagenz **67**, das nur unfunktionalisierte PEG-*Spacer* aufweist, wurde auch die Reaktion mit dem Hydroxylaminreagenz **68**, das neben zwei PEG-*Spacer*n auch einen biotinylierten *Spacer* trägt, untersucht. Hier konnte ebenfalls vollständige Umsetzung des Startmaterials unter Bildung der gewünschten Oximspezies bei etwas längerer Reaktionsdauer unter analogen Bedingungen detektiert werden.

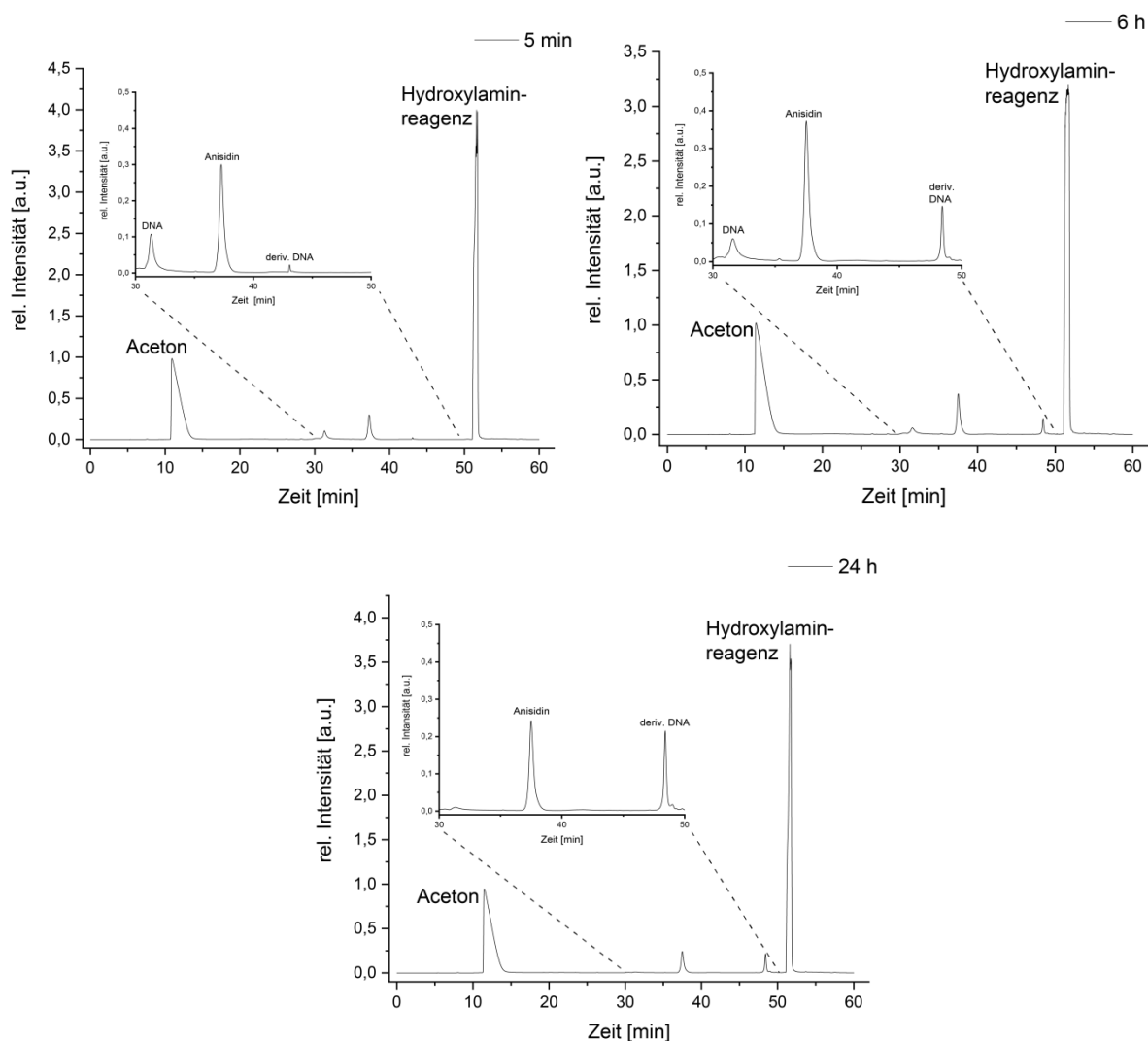


Abbildung 6-30: HPL-Chromatogramme der Oligonukleotid-Derivatisierungsexperimente nach 5 min, 6 h und 24 h; ausgewählter Bereich hervorgehoben.

6.4.3. Zusammenfassung und Ausblick

Im Rahmen dieses Projekts wurden verschiedene, sterisch anspruchsvolle Hydroxylaminreagenzien synthetisiert. Diese sollten zur Derivatisierung von fdC auf DNA-Ebene eingesetzt werden, um bei Echtzeitsequenzierexperimenten möglicherweise das derivatisierte fdC anhand des großen Substituenten selektiv erfassen zu können. Aus diesem Grund wurden Hydroxylaminreagenzien verschiedenen sterischen Anspruchs synthetisiert. Dabei wurden auch Reagenzien hergestellt, die einen abspaltbaren Biotinrest zur selektiven Anreicherung der derivatisierten Verbindungen vor den Sequenzierexperimenten ermöglichen könnten. Untersuchungen des Reaktionsverhaltens einiger exemplarischer Derivatisierungsreagenzien mit einem fdC-enhalten Oligonukleotid konnte optimierte Reaktionsbedingungen liefern. Erste Sequenzierexperimente mit einigen der hier vorgestellten

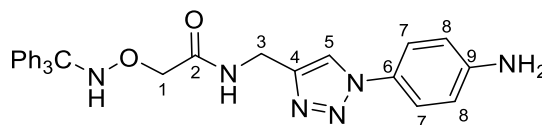
Hydroxylaminreagenzien werden in der Arbeitsgruppe von *Prof. Jörn Walter* an der *Universität des Saarlandes* durchgeführt.

6.4.4. Projektbeiträge

Das Design und die Synthesen aller der in diesem Kapitel synthetisierten, sterisch anspruchsvollen Hydroxylaminreagenzien wurden von mir durchgeführt. Das dreifach alkinylierte Gallussäure-Derivat wurde von *Bastien Viverge* synthetisiert und mir bereitgestellt. Die Bindungsstudien der Verbindungen mit einem fdC-enthaltenden DNA-Oligonukleotid wurden von mir durchgeführt. Das Design dieser Experimente beruht auf von *Dr. Matthias Kurz* durchgeführten Studien zur Reaktion des bekannten Hydroxylamin-Reagenzes mit fdC-enthaltenden Oligonukleotiden. Das in meinen Bindungsstudien verwendete fdC-enthaltende DNA-Oligonukleotid wurde von *Dr. Edris Parsa* synthetisiert. Erste Sequenzierexperimente, die eine Derivatisierung von fdC mit den hier vorgestellten Derivatisierungsreagenzien untersuchen, werden im Arbeitskreis von *Prof. Jörn Walter* an der *Universität des Saarlandes* durchgeführt.

6.4.5. Experimentalteil

***N*-((1-(4-Aminophenyl)-1H-1,2,3-triazol-4-yl)methyl)-2-((tritylamino)oxy)acetamid (42)**



Das Trityl-geschütztes Alkin^[53,124] **41** (326 mg, 0.88 mmol, 1.0 eq.) wurde in 67 mL gründlich entgastem THF (*freeze-pump-thaw*, 3x) gelöst und unter Argon-Atmosphäre mit 150 mg 4-Azidoanilin-Hydrochlorid (0.88 mmol, 1.0 eq.), 460 μ L DIPEA (2.64 mmol, 3.0 eq.) und schließlich mit 334 mg CuI (1.76 mmol, 2.0 eq.) versetzt. Das Reaktionsgemisch wurde 23 h bei Raumtemperatur verrührt und anschließend *in vacuo* zur Trockene eingengt. Der Rückstand wurde über Säulenchromatographie an Silicagel (CH₂Cl₂:MeOH 25:1 \rightarrow CH₂Cl₂:MeOH 20:1) aufgereinigt, um 387 mg des Cycloadditionsproduktes **42** (0.77 mmol, 87%) als gelben Feststoff zu liefern.

¹H-NMR (600 MHz, CDCl₃) δ /ppm = 7.70 (s, 1H, 5-H), 7.45 – 7.40 (m, 2H, 7-H), 7.34 – 7.21 (m, 15H, CPh₃), 6.77 – 6.73 (m, 2H, 8-H), 6.69 (s, 1H, ONH), 6.31 (t, J = 5.4 Hz, 1H, (C=O)NH), 4.41 (d, J = 6.0 Hz, 2H, 3-H), 4.27 (s, 2H, 1-H), 3.88 (s, 2H, NH₂).

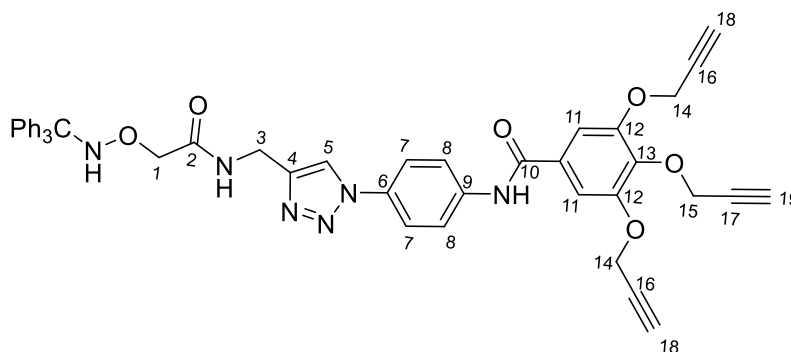
¹³C-NMR (201 MHz, CDCl₃) δ /ppm = 169.61 (2-C), 147.23 (9-C), 144.88 (4-C), 143.78 (C_{Ar}), 128.98 (C_{Ar-CPh3}), 128.71 (6-C), 128.11 (C_{Ar-CPh3}), 127.32 (C_{Ar-CPh3}), 122.36 (7-C), 120.66 (5-C), 115.43 (8-C), 74.47 (CPh₃), 73.59 (1-C), 34.61 (3-C).

HRMS (ESI⁺): m/z berechnet für C₃₀H₂₉N₆O₂⁺ [M+H]⁺ 505.2347; gefunden: 505.2348.

IR (ATR): ν (cm⁻¹) = 3370 (br, m), 2918 (s), 2851 (s), 2356 (w), 1736 (w), 1657 (w), 1522 (w), 1469 (w), 1266 (w), 1177 (w), 1059 (w), 739 (m), 703 (s).

R_f (CH₂Cl₂:MeOH 20:1): 0.29.

3,4,5-Tris(prop-2-yn-1-yloxy)-N-(4-(4-((2-((tritylamino)oxy)acetamido)methyl)-1H-1,2,3-triazol-1-yl)phenyl)benzamid (44)



Das Cycloadditionsprodukt **42** (219 mg, 0.43 mmol, 1.0 eq.) und das dreifach alkinylierte Gallussäure-Derivat^[126] **43** (123 mg, 0.43 mmol, 1.0 eq.) wurden in 3.6 mL trockenem CH₂Cl₂ suspendiert und mit 91 µL DIPEA (0.52 mmol, 1.2 eq.) versetzt, woraufhin eine klare Lösung erhalten wurde. Dem Gemisch wurde HBTU (198 mg, 0.52 mmol, 1.2 eq) zugefügt und die Lösung wurde bei Raumtemperatur für 20 h verrührt. Auf Grund einer unvollständigen Reaktion wurden weitere 123 mg **43** (0.43 mmol, 1.0 eq.), 91 µL DIPEA (0.52 mmol, 1.2 eq.) und 198 mg HBTU (0.52 mmol, 1.2 eq) zugesetzt und es wurde weitere 23 h bei Raumtemperatur gerührt. Daraufhin wurde das Reaktionsgemisch in gesättigte wässrige NH₄Cl-Lösung gegossen (150 mL) und mit EtOAc extrahiert (3 x 150 mL). Die vereinigten organischen Phasen wurden mit gesättigter wässriger NaHCO₃-Lösung (150 mL) und gesättigter wässriger Natriumchloridlösung (150 mL) gewaschen, über Na₂SO₄ getrocknet und *in vacuo* zur Trockene eingeeengt. Aufreinigung des Rückstandes über Säulenchromatographie an Silicagel (iHex:EtOAc 1:1 → iHex:EtOAc 1:2 → iHex:EtOAc 0:1) lieferte 156 mg des Kupplungsproduktes **44** (0,20 mmol, 47%) als schmutzig-weißen Feststoff.

¹H-NMR (600 MHz, CDCl₃) δ/ppm = 8.17 (s, 1H, (C=O)-NH-C_{Ar}), 7.82 – 7.78 (m, 3H, 5-H, 7-H), 7.69 – 7.63 (m, 2H, 8-H), 7.33 – 7.18 (m, 17H, 11-H, CPh₃), 6.76 (s, 1H, , ONH), 6.40 (t, *J* = 6.0 Hz, 1H, (C=O)NH-C_{Alk}), 4.84 – 4.79 (m, 6H, 14-H, 15-H), 4.40 (t, *J* = 6.0 Hz, 2H, 3-H), 4.26 (s, 2H, 1-H), 2.57 – 2.46 (m, 3H, 18-H, 19-H).

¹³C-NMR (201 MHz, CDCl₃) δ/ppm = 169.84 (C-2), 168.49 (C-10), 165.17 ((C=O)-NH-C_{Ar}), 151.85 (12-C), 145.35 (4-C), 143.76 (C_{Ar}-CPh₃), 140.71 (C_{Ar}-CPh₃), 140.52 (13-C), 138.56 (9-C), 133.23 (6-C), 130.44 (11-C), 129.27 (C_{Ar}-CPh₃), 128.98 (C_{Ar}-CPh₃),

128.11 ($\text{C}_{\text{Ar-CPh}_3}$), 127.34 ($\text{C}_{\text{Ar-CPh}_3}$), 121.29 (8-C), 121.19 (7-C), 120.50 (5-C), 108.33 ($\text{C}_{\text{Ar-CPh}_3}$), 108.25 ($\text{C}_{\text{Ar-CPh}_3}$), 78.76 (C-17), 78.12 (C-16), 75.91 (C-18), 74.49 (C-19), 73.58 (1-C), 60.56 (C-15), 57.53 (C-14), 34.58 (3-C).

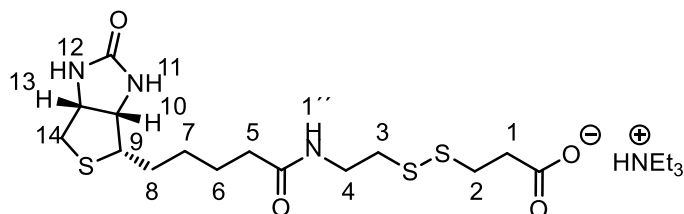
HRMS (ESI⁺): m/z berechnet für $\text{C}_{46}\text{H}_{39}\text{N}_6\text{O}_6^+$ $[\text{M}+\text{H}]^+$ 771.2926; gefunden: 771.2912.

HRMS (ESI⁻): m/z berechnet für $\text{C}_{46}\text{H}_{37}\text{N}_6\text{O}_6^-$ $[\text{M}-\text{H}]^-$ 769.2780; gefunden: 769.2800.

IR (ATR): ν (cm^{-1}) = 3394 (w), 3261 (w), 3230 (w), 1665 (m), 1647 (s), 1607 (w), 1554 (m), 1520 (s), 1497 (m), 1382 (w), 1336 (s), 1249 (w), 1194 (w), 1106 (s), 1039 (m), 988 (s), 840 (s), 767 (m), 742 (s), 711 (s), 700 (s).

R_f (iHex: EtOAc 1:2): 0.38.

Triethylammonium-3-((2-(+)-biotinamido)ethyl)disulfanyl)propanoat (**50**)



Unter Argon-Atmosphäre wurden 404 mg der Aminocarbonsäure^[128] **49** (2.23 mmol, 1.0 eq.) in 11.0 mL trockenem DMF suspendiert, mit 761 mg Biotin-NHS-Ester^[129] (2.23 mmol, 1.0 eq.) versetzt und 1.40 mL NEt_3 (10.03 mmol, 4.5 eq.) zugetropft, worauf sich eine klare Lösung bildete. Das Reaktionsgemisch wurde 16 h bei Raumtemperatur verrührt, wobei sich im weiteren Verlauf eine gelbe Suspension entwickelte. Das Gemisch wurde *in vacuo* bei 50 °C zur Trockene eingeeengt und der Rückstand durch Säulenchromatographie an Silicagel (CH_2Cl_2 :MeOH 15:1 → CH_2Cl_2 :MeOH 10:1 → CH_2Cl_2 :MeOH 5:1, jew. mit 0.25% HOAc) aufgereinigt. Coevaporation des Produkts mit Toluol (3 x 30 mL) lieferte 815 mg des Triethylammoniumsalzes des Kupplungsprodukts **50** (1,60 mmol, 72%) als weißen Feststoff.

¹H-NMR (400 MHz, $\text{DMSO}-d_6$) δ /ppm = 8.04 (t, J = 5.6 Hz, 1H, 1'-H), 6.44 (s, 1H, 11-H), 6.37 (s, 1H, 12-H), 4.30 (dd, J = 7.8, 5.1 Hz, 1H, 13-H), 4.13 (ddd, J = 7.7, 4.5, 1.8 Hz, 1H, 10-H), 3.35 – 3.27 (m, 2H, 4-H), 3.09 (ddd, J = 8.4, 6.1, 4.4 Hz, 1H, 9-H), 2.89 (q, J = 7.2 Hz, 6H, $\text{HN}(\text{CH}_2\text{CH}_3)_3$), 2.85 – 2.72 (m, 4H, 2-H, 3-H), 2.61 – 2.54

(m, 2H, 1-H), 2.06 (t, $J = 7.4$ Hz, 2H, 5-H), 1.65 – 1.23 (m, 6H, 6-H, 7-H, 8-H), 1.14 (t, $J = 7.3$ Hz, 9H, $\text{HN}(\text{CH}_2\text{CH}_3)_3$).

^{13}C -NMR (101 MHz, $\text{DMSO}-d_6$) $\delta/\text{ppm} = 172.16$ (CO-NH-CH₂), 162.70 (NH-CO-NH), 61.03 (10-C), 59.20 (13-C), 55.38 (9-C), 45.30 ($\text{HN}(\text{CH}_2\text{CH}_3)_3$), 39.83 (2/3-C), 37.87 (4-C), 37.34 (2/3-C), 35.12 (5-C), 28.15 (6/7/8-C), 28.01 (6/7/8-C), 25.21 (6/7/8-C), 9.09 ($\text{HN}(\text{CH}_2\text{CH}_3)_3$).

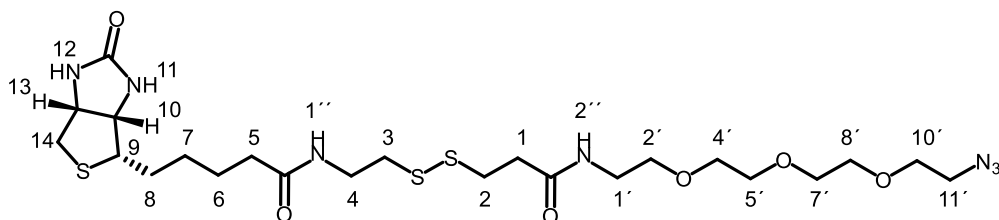
HRMS (ESI⁺): m/z berechnet für $\text{C}_{15}\text{H}_{26}\text{N}_3\text{O}_4\text{S}_3^+$ $[\text{M}+\text{H}]^+$ 408.1080; gefunden: 408.1083.

HRMS (ESI⁻): m/z berechnet für $\text{C}_{15}\text{H}_{24}\text{N}_3\text{O}_4\text{S}_3^-$ $[\text{M}-\text{H}]^-$ 406.0934; gefunden: 406.0928.

IR (ATR): ν (cm^{-1}) = 2978 (m), 2946 (m), 2739 (w), 2603 (s), 2497 (s), 1704 (m), 1658 (w), 1642 (w), 1476 (s), 1444 (m), 1398 (s), 1384 (w), 1172 (m), 1072 (w), 1036 (s), 851 (w), 808 (m).

R_f (CH_2Cl_2 :MeOH 5:1 + 0.25% HOAc): 0.44.

3-((2-(+)-Biotinamido)ethyl)disulfanyl)propansäure-(11'-azido-3',6',9'-trioxaundecan-1'-amid) (51)



Eine Suspension von 500 mg biotinylierter Carbonsäure **50** (0.98 mmol, 1.0 eq.) in 6.0 mL trockenem DMF wurde tropfenweise mit 11-Azido-3,6,9-trioxaundecan-1-amin (234 μL , 1.18 mmol, 1.2 eq) und DIPEA (205 μL , 1.18 mmol, 1.2 eq.) versetzt. Daraufhin wurden 448 mg HATU (1.18 mmol, 1.2 eq.) zugesetzt und das Gemisch wurde 20 h bei Raumtemperatur verrührt. Anschließend wurden weitere 117 μL 11-Azido-3,6,9-trioxaundecan-1-amin (0.59 mmol, 0.6 eq), 103 μL DIPEA (0.59 mmol, 0.6 eq) und 224 mg HATU (0.59 mmol, 0.6 eq) zugegeben und das Reaktionsgemisch weitere 3 h bei Raumtemperatur verrührt. Die erhaltene Lösung wurde dann in gesättigte wässrige NaHCO_3 -Lösung (100 mL) gegossen und dreimal mit CH_2Cl_2 (3 x 100 mL) extrahiert. Die vereinigten organischen Phasen wurden über

Na₂SO₄ getrocknet, *in vacuo* zur Trockene eingengt (50 °C) und schließlich wurde der erhaltene Rückstand mittels Säulenchromatographie an Silicagel (CH₂Cl₂:MeOH 20:1 → CH₂Cl₂:MeOH 15:1 → CH₂Cl₂:MeOH 10:1 → CH₂Cl₂:MeOH 8:1) aufgereinigt, um 390 mg des Kupplungsprodukts **51** (0.64 mmol, 65%) als schmutzig-weißen Feststoff zu liefern.

¹H-NMR (600 MHz, CDCl₃) δ/ppm = 6.98 (s, 1H, 1''-H), 6.81 (t, *J* = 5.6 Hz, 1H, 2''-H), 6.41 (s, 1H, 11-H), 5.58 (s, 1H, 12-H), 4.51 (dd, *J* = 7.8, 4.9 Hz, 1H, 13-H), 4.34 – 4.29 (m, 1H, 10-H), 3.69 – 3.61 (m, 10H, 4'-H, 5'-H, 7'-H, 8'-H, 10'-H), 3.60 – 3.56 (m, 2H, 2'-H), 3.54 (qt, *J* = 6.2, 3.3 Hz, 2H, 4-H), 3.45 (qd, *J* = 5.2, 2.1 Hz, 2H, 1'-H), 3.39 (dd, *J* = 5.5, 4.6 Hz, 2H, 11'-H), 3.14 (td, *J* = 7.4, 4.5 Hz, 1H, 9-H), 3.02 – 2.94 (m, 2H, 2-H), 2.90 (dd, *J* = 12.8, 4.9 Hz, 1H, 14-H), 2.84 (t, *J* = 6.3 Hz, 2H, 3-H), 2.73 (d, *J* = 12.8 Hz, 1H, 14-H), 2.62 (td, *J* = 7.2, 4.7 Hz, 2H, 1-H), 2.28 – 2.19 (m, 2H, 5-H), 1.80 – 1.60 (m, 4H, 6-H, 8-H), 1.45 (p, *J* = 7.4 Hz, 2H, 7-H).

¹³C-NMR (151 MHz, CDCl₃) δ/ppm = 173.69 ((CH₂)₄-CO-NH), 171.48 ((CH₂)₂-CO-NH), 164.07 (NH-CO-NH), 70.80 - 70.16 (5 x s, 4'-C, 5'-C, 7'-C, 8'-C, 10'-C), 69.85 (2'-C), 61.93 (10-C), 60.30 (13-C), 55.84 (9-C), 50.81 (11'-C), 40.75 (14-C), 39.54 (1'-C), 38.58 (3-C), 38.44 (4-C), 36.32 (1-C), 35.81 (5-C), 34.12 (2-C), 28.21 (7-C), 28.15 (8-C), 25.71 (6-C).

HRMS (ESI⁺): *m/z* berechnet für C₂₃H₄₂N₇O₆S₃⁺ [M+H]⁺ 608.2353; gefunden: 608.2350.

IR (ATR): ν (cm⁻¹) = 3284 (m), 2922 (m), 2102 (m), 1693 (s), 1640 (s), 1546 (s), 1462 (m), 1351 (w), 1265 (s), 1102 (s), 1028 (s), 940 (w), 855 (w).

R_f (CH₂Cl₂:MeOH 5:1): 0.62.

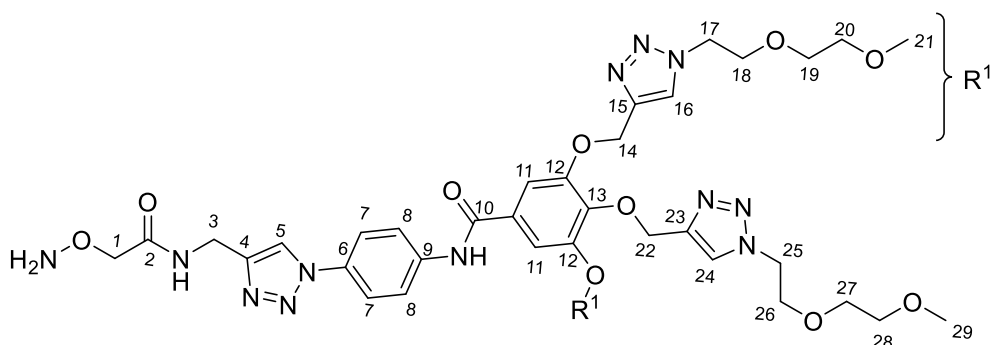
Generelle Vorschrift zur CuAAC von **44 mit den Azido-Spacern **47** und/oder **51****

Die dreifach alkinylierte Verbindung **44** (1.0 eq.) wurde zusammen mit den gewünschten *Spacern* (3.1 eq.; entweder 3.1 eq. eines unfunktionalisierten *Spacers* **44** oder 3.1 eq. eines 2:1 Stoffmengengemisches aus unfunktionalisierten **44** und biotinylierten *Spacern* **51**) in einem gründlich entgasten (*freeze-pump-thaw*, 3 x) Gemisch aus THF und ddH₂O gelöst (1:1, 46 mL/mmol). Anschließend wurde der Kupfer(I)-bromid-Dimethylsulfidkomplex (0.3 eq.) zugegeben und das Reaktionsgemisch wurde 72 h bei Raumtemperatur verrührt. Anschließend wurde das Gemisch

in vacuo zur Trockene eingengt. Der Rückstand wurde in einem Gemisch aus CH₂Cl₂ und wässriger 6 M Salzsäure aufgenommen (1:1, 30 mL/mmol) und 1 h bei Raumtemperatur verrührt. Anschließend wurden die Phasen getrennt und die wässrige Phase wurde mit CH₂Cl₂ gewaschen (5 x 75 mL/mmol). Die wässrige Phase wurde anschließend mit konz. wässrigen Ammoniak neutralisiert und das Gemisch wurde über Gefriertrocknung getrocknet. Der Rückstand wurde in ddH₂O aufgenommen, über einen Spritzenvorsatzfilter gefiltert und mittels *reversed-phase* HPLC (0% → 80% MeCN in ddH₂O in 45 min) aufgereinigt, um die entschützten CuAAC-Reaktionsprodukte zu erhalten.

Im Folgenden werden eine vollständig unfunktionalisierte Verbindung **67** und eine einfach biotinylierte Verbindung **68** charakterisiert:

***N*-(4-(4-((2-(Aminoxy)acetamido)methyl)-1H-1,2,3-triazol-1-yl)phenyl)-3,4,5-tris-((1-(2-(2-methoxyethoxy)ethyl)-1H-1,2,3-triazol-4-yl)methoxy)benzamid (67)**



Ausbeute: 82% über 2 Stufen, weißer Feststoff.

¹H-NMR (600 MHz, D₂O) δ/ppm = 8.04 (s, 1H, 24-H), 7.98 (s, 2H, 16-H), 7.70 (d, *J* = 2.1 Hz, 1H, 5-H), 7.48 (d, *J* = 8.0 Hz, 2H, 7/8-H), 7.30 (d, *J* = 8.3 Hz, 2H, 7/8-H), 7.03 (s, 2H, 11-H), 4.92 (s, 4H, 14-H), 4.84 – 4.81 (m, 2H, 22-H), 4.53 – 4.50 (m, 4H, 19/20-H), 4.46 – 4.36 (m, 4H, 3-H, 27/28-H), 4.24 – 4.17 (m, 2H, 27/28-H), 3.88 – 3.81 (m, 4H, 19/20-H), 3.75 (d, *J* = 4.9 Hz, 2H, 26-H), 3.50 – 3.35 (m, 12H, 17/18/25/26-H), 3.17 – 3.15 (m, 9H, 21-H, 29-H).

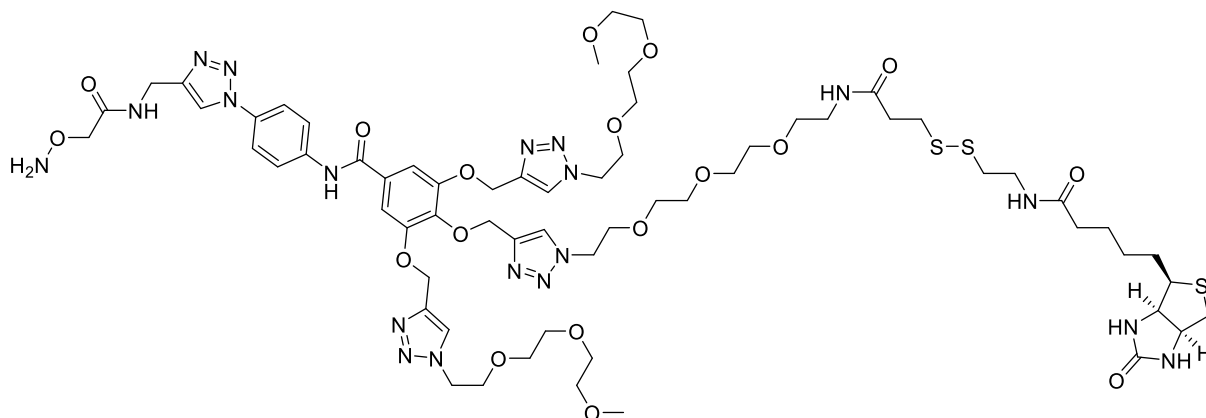
¹³C-NMR (151 MHz, D₂O) δ/ppm = 172.21 ((C=O)-NH), 165.48 ((C=O)-NH), 151.18, 144.85, 143.19, 142.82, 139.07, 138.03, 131.98, 129.03, 125.42 (5-C), 125.28 (16-C), 121.43 (7/8-C), 121.10 (24-C), 120.12 (7/8-C), 106.77 (11-C), 73.71 (27/28-C),

70.79 - 69.36 (4 x s, 17/18/25/26-C), 68.61 (19/20-C), 65.02 (3-C), 61.81 (14-C), 57.88 (21/29-C), 49.96, 49.82, 33.80 (3-C).

HRMS (ESI⁺): m/z berechnet für C₄₂H₅₈N₁₅O₁₂⁺ [M+H]⁺ 964.4384; gefunden: 964.4374.

HRMS (ESI⁻): m/z berechnet für C₄₂H₅₆N₁₅O₁₂⁻ [M-H]⁻ 962.4238; gefunden: 962.4241.

N-(4-(4-((2-(Aminooxy)acetamido)methyl)-1H-1,2,3-triazol-1-yl)phenyl)-4-((1-(13,21-dioxo-25-((3aR,4R,6aS)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)-3,6,9-trioxa-16,17-dithia-12,20-diazapentacosyl)-1H-1,2,3-triazol-4-yl)methoxy)-3,5-bis((1-(2-(2-(2-methoxyethoxy)ethoxy)ethyl)-1H-1,2,3-triazol-4-yl)methoxy)-benzamid (68)



Ausbeute: 4% über 2 Stufen, weißer Feststoff.

¹H-NMR (800 MHz, D₂O) δ/ppm = 8.29 (s, 1H, C-H_{Triazol}), 8.14 (s, 2H, C-H_{Triazol}), 7.86 (s, 1H, C-H_{Triazol}), 7.74 (d, *J* = 8.3 Hz, 2H, C-H_{Ar}), 7.65 (d, *J* = 8.4 Hz, 2H, C-H_{Ar}), 7.33 (s, 2H, C-H_{Ar}), 5.26 – 5.15 (m, 4H, X-CH_{2,Aliph}), 5.10 (d, *J* = 5.1 Hz, 2H, X-CH_{2,Aliph}), 4.66 – 4.46 (m, 10H, X-CH_{2,Aliph}), 4.33 – 4.23 (m, 3H, X-CH_{2,Aliph}), 3.99 – 3.89 (m, 5H, X-CH_{2,Aliph}), 3.84 (q, *J* = 6.4, 5.8 Hz, 2H, X-CH_{2,Aliph}), 3.63 – 3.40 (m, 25H, X-CH_{2,Aliph}), 3.36 – 3.19 (m, 12H, X-CH_{2,Aliph} OCH₃), 3.14 – 3.04 (m, 1H, X-CH_{2,Aliph}), 2.88 – 2.47 (m, 9H, X-CH_{2,Aliph}, CH_{2,Aliph}), 2.15 – 2.05 (m, 2H, CH_{2,Aliph}), 1.63 – 1.15 (m, 6H, CH_{2,Aliph}). [X=O,N,S]

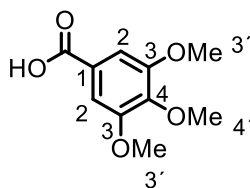
¹³C-NMR (201 MHz, D₂O) δ/ppm = 176.31 ((C=O)-NH), 173.76 ((C=O)-NH), 172.40 ((C=O)-NH), 165.06 ((C=O)-NH), 151.53 ((C=O)-NH), 145.13 (C_{Ar}), 142.85 (C_{Ar}), 139.24 (C_{Ar}), 138.27 (C_{Ar}), 132.67 (C_{Ar}), 129.93 (C_{Ar}), 129.58 (C_{Ar}), 125.60 (C_{Ar}),

125.59 (**C**_{Ar}), 122.20 (**C**_{Ar}), 122.18 (**C**_{Ar}), 121.01 (**C**_{Ar}), 121.00 (**C**_{Ar}), 115.28 (**C**_{Ar}), 107.25 (**C**_{Ar}), 73.74 (X-**CH**_{Aliph}), 70.92 (X-**CH**_{Aliph}), 69.66 (X-**CH**_{Aliph}), 69.45 (X-**CH**_{Aliph}), 69.37 (X-**CH**_{Aliph}), 68.73 (X-**CH**_{Aliph}), 65.12 (X-**CH**_{Aliph}), 62.04 (X-**CH**_{Aliph}), 61.94 (X-**CH**_{Aliph}), 60.12 (X-**CH**_{Aliph}), 58.05 (X-**CH**_{Aliph}), 57.96 (X-**CH**_{Aliph}), 55.33 (X-**CH**_{Aliph}), 50.19 (X-**CH**_{Aliph}), 50.13 (X-**CH**_{Aliph}), 39.69 (X-**CH**_{Aliph}), 38.99 (X-**CH**_{Aliph}), 37.79 (X-**CH**_{Aliph}), 36.70 (X-**CH**_{Aliph}), 35.37 (**CH**_{Aliph}), 34.93 (**CH**_{Aliph}), 33.96 (**CH**_{Aliph}), 33.30 (**CH**_{Aliph}), 27.91 (**CH**_{Aliph}), 27.90 (**CH**_{Aliph}), 27.62 (**CH**_{Aliph}), 25.11 (**CH**_{Aliph}).
[X=O,N,S]

HRMS (ESI⁺): m/z berechnet für C₆₄H₉₆N₁₉O₁₈S₃⁺ [M+H]⁺ 1514.6337; gefunden: 1514.6307; m/z berechnet für C₆₄H₉₇N₁₉O₁₈S₃⁺ [M+2H]²⁺ 757.8205; gefunden 757.8192.

HRMS (ESI⁻): m/z berechnet für C₆₄H₉₄N₁₉O₁₈S₃⁻ [M-H]⁻ 1512.6192; gefunden: 1512.6237.

3,4,5-Trimethoxybenzoessäure^[131] (**55**)

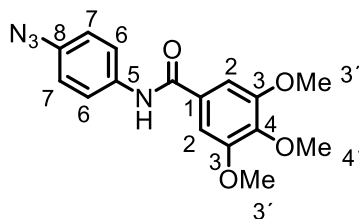


200 mg Methyl-3,4,5-trimethoxybenzoat **54** (0.88 mmol, 1.0 eq.) wurden in 8.3 mL MeOH gelöst und tropfenweise mit einer Lösung von 185 mg Lithiumhydroxid-Monohydrat (4.40 mmol, 5.0 eq.) in 2.2 mL ddH₂O versetzt. Daraufhin wurde das Gemisch 22 h bei Raumtemperatur verrührt, anschließend *in vacuo* zur Trockene eingengt, der Rückstand in ddH₂O (100 mL) aufgenommen und mit wässriger Salzsäure (1 M) angesäuert. Dann wurde mit EtOAc extrahiert (2 x 100 mL), die vereinigten organischen Phasen wurden mit gesättigter wässriger Natriumchloridlösung (100 mL) gewaschen, über Na₂SO₄ getrocknet und Lösungsmittel wurden *in vacuo* entfernt. Dabei wurden 183 mg 3,4,5-Trimethoxybenzoessäure **55** (0.86 mmol, 98%) als weißer Feststoff erhalten, der ohne weitere Aufreinigung direkt für die nächste Stufe verwendet werden konnte.

¹H-NMR (400 MHz, CDCl₃) δ/ppm = 7.38 (s, 2H, 2-H), 3.94 (s, 3H, 4'-H), 3.93 (s, 6H, 3'-H)

HRMS (ESI⁻): m/z berechnet für C₁₆H₁₅N₄O₄⁻ [M-H]⁻ 211.0612; gefunden: 211.0612.

N-(4-Azidophenyl)-3,4,5-trimethoxybenzamid (56)



56 mg des Benzoessäure-Derivats **55** (0.26 mmol, 1.0 eq.) wurden in 1.5 mL trockenem CH₂Cl₂ gelöst und mit DIPEA (55 µL, 0.32 mmol, 1.2 eq.), 4-Azidoanilin-Hydrochlorid (45 mg, 0.26 mmol, 1.0 eq.) und HBTU (120 mg, 0.32 mmol, 1.2 eq.) versetzt. Das Reaktionsgemisch wurde daraufhin 23 h bei Raumtemperatur verrührt, anschließend in gesättigte wässrige NH₄Cl-Lösung (50 mL) gegossen, und die wässrige Phase wurde mit EtOAc (100 mL) extrahiert. Die organische Phase wurde dann mit gesättigter wässriger NaHCO₃-Lösung (50 mL) und gesättigter wässriger Natriumchloridlösung (50 mL) gewaschen, über Na₂SO₄ getrocknet und *in vacuo* zur Trockene eingeeengt, Aufreinigung des Rohprodukts mittels Säulenchromatographie an Silicagel (CH₂Cl₂:MeOH 20:1 → CH₂Cl₂:MeOH 10:1) lieferte das Kupplungsprodukt **56** als rotbraunen Feststoff (74 mg, 0.23 mmol, 86%).

¹H-NMR (400 MHz, CDCl₃) δ/ppm = 7.76 (s, 1H, NH), 7.67 – 7.60 (m, 2H, 6-H), 7.06 (s, 2H, 2-H), 7.05 – 7.02 (m, 2H, 7-H), 3.93 (s, 6H, 3'-H), 3.91 (s, 3H, 4'-H).

¹³C-NMR (101 MHz, CDCl₃) δ/ppm = 165.58 (CO-NH), 153.51 (3-C), 141.45 (4-C), 136.24 (8-C), 134.98 (5-C), 130.29 (1-C), 121.87 (6-C), 119.74 (7-C), 104.56 (2-C), 61.13 (4'-C), 56.56 (3'-C).

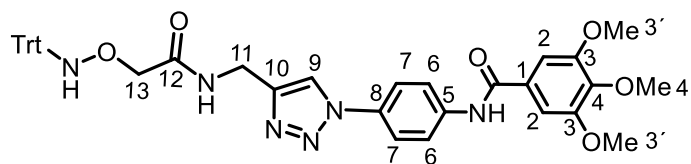
HRMS (ESI⁺): m/z berechnet für C₁₆H₁₇N₄O₄⁺ [M+H]⁺ 329.1244; gefunden: 329.1247.

HRMS (ESI⁻): m/z berechnet für C₁₆H₁₅N₄O₄⁻ [M-H]⁻ 327.1099; gefunden: 327.1100.

IR (ATR): ν (cm⁻¹) = 3310 (w), 2940 (w), 2839 (w), 2361 (w), 2117 (s), 1789 (w), 1647 (m), 1584 (s), 1533 (m), 1505 (s), 1464 (w), 1415 (m), 1335 (s), 1290 (w), 1238 (m), 1175 (w), 1127 (s), 1003 (w), 835 (w), 734 (w).

R_f (CH₂Cl₂:MeOH 20:1): 0.57.

3,4,5-Trimethoxy-*N*-(4-(4-((2-((tritylamino)oxy)acetamido)methyl)-1H-1,2,3-triazol-1-yl)phenyl)benzamid (69)



In einem Schenk-Kolben wurde ein Gemisch aus 1.8 mL THF und 1.8 mL ddH₂O gründlich entgast (*freeze-pump-thaw*, 3 x). Anschließend wurden 63 mg Azid **56** (0.19 mmol, 1.0 eq.), 78 mg Trityl-geschütztes Alkin^[53,124] **41** (0.21 mmol, 1.1 eq.) und schließlich 12 mg Kupfer(I)-bromid-Dimethylsulfidkomplex (58 µmol, 0.3 eq.) zugegeben und das Gemisch wurde 23 h bei Raumtemperatur verrührt. Das Reaktionsgemisch wurde *in vacuo* zur Trockene eingeeengt und direkt mittels Säulenchromatographie an Silicagel (CH₂Cl₂:MeOH 25:1 → CH₂Cl₂:MeOH 15:1) aufgereinigt, um das Cycloadditionsprodukt **69** (133 mg, 0.19 mmol, quant.) als leicht gelblichen Feststoff zu erhalten.

¹H-NMR (400 MHz, DMSO-*d*₆) δ/ppm = 10.36 (s, 1H, ONH-CPh₃), 8.55 (s, 1H, 9-H), 8.30 – 8.20 (m, 2H, (C=O)-NH-CH₂, (C=O)-NH-CH_{Ar}), 8.00 – 7.80 (m, 4H, 6-H, 7-H), 7.41 – 7.18 (m, 17H, 2-H, CPh₃), 4.45 (d, *J* = 5.7 Hz, 2H, 11-H), 3.88 (s, 6H, 3'-H), 3.86 (s, 2H, 13-H), 3.74 (s, 3H, 4'-H).

¹³C-NMR (101 MHz, DMSO-*d*₆) δ/ppm = 169.66 ((C=O)-NH-CH₂), 165.13 ((C=O)-NH-CH_{Ar}), 152.67 (3'-C), 145.93 (10-C), 144.16 (1-C), 140.43 (4'-C), 139.28 (8-C), 132.18 (5-C), 129.76 (CPh₃), 128.92 (CPh₃), 127.67 (CPh₃), 126.81 (CPh₃), 121.31 (C-6/7), 120.93 (C-9), 120.37 (C-6/7), 105.34 (C-2), 73.72 (13-C), 60.17 (C-4'), 56.14 (C-3'), 33.85 (C-11).

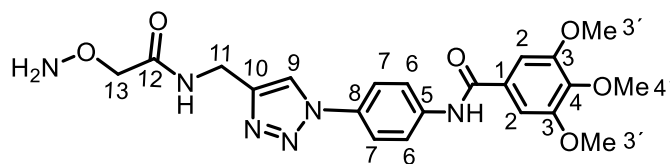
HRMS (ESI⁺): *m/z* berechnet für C₄₀H₃₉N₆O₆⁺ [M+H]⁺ 699.2926; gefunden: 699.2942.

HRMS (ESI⁻): *m/z* berechnet für C₄₀H₃₇N₆O₆⁻ [M-H]⁻ 697.2780; gefunden: 697.2768.

IR (ATR): ν (cm⁻¹) = 3292 (w), 3234 (w), 2364 (w), 1648 (m), 1522 (s), 1417 (m), 1338 (s), 1239 (m), 1124 (s), 1004 (w), 840 (m), 750 (s), 708 (s).

R_f (CH₂Cl₂:MeOH 10:1): 0.54.

N-(4-(4-((2-(Aminooxy)acetamido)methyl)-1H-1,2,3-triazol-1-yl)phenyl)-3,4,5-trimethoxybenzamid (57)



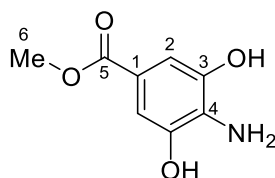
49 mg der Trityl-geschützten Verbindung **69** (70 μmol , 1.0 eq.) wurden in 1.5 mL CH_2Cl_2 suspendiert und mit 1.5 mL wässriger Salzsäure (6 M) versetzt. Das Gemisch wurde daraufhin 1 h bei Raumtemperatur verrührt, die Phasen wurden getrennt, die wässrige Phase wurde mit CH_2Cl_2 gewaschen (5 x 5 mL) und anschließend mit konzentriertem wässrigem Ammoniak (25%) neutralisiert. Nach Entfernen des Lösungsmittels durch Gefriertrocknung wurde der Rückstand in 5 mL ddH₂O aufgenommen, über einen Spritzenvorsatzfilter gefiltert und danach mittels *reversed-phase* HPLC (0% \rightarrow 80% MeCN in ddH₂O in 45 min) aufgereinigt, um die entschützte Verbindung **57** in Form eines farblosen Feststoffs (2.1 mg, 5 μmol , 7%) zu liefern.

¹H-NMR (400 MHz, D₂O) δ/ppm = 8.24 (s, 1H, 9-H), 7.69 – 7.61 (m, 4H, 6-H, 7-H), 7.14 (s, 2H, 2-H), 4.52 (s, 2H, 11-H), 4.15 (s, 2H, 13-H), 3.83 (s, 6H, 3'-H), 3.74 (s, 3H, 4'-H).

HRMS (ESI⁺): m/z berechnet für C₂₁H₂₅N₆O₆⁺ [M+H]⁺ 457.1830; gefunden: 457.1830.

HRMS (ESI⁻): m/z berechnet für C₂₁H₂₃N₆O₆⁻ [M-H]⁻ 455.1685; gefunden: 455.1683.

4-Amino-3,5-dihydroxybenzoesäuremethylester (59)



Gallussäure **58** (1.00 g, 5.88 mmol, 1.0 eq.) wurde in 15 mL konzentriertem wässrigen Ammoniak gelöst und ohne Schutzgasatmosphäre im offenen Reaktionsgefäß bei Raumtemperatur für 15 min verrührt. Anschließend wurden weitere 20 mL konzentrierter wässriger Ammoniak zugesetzt und das Gemisch für weitere 15 min gerührt. Daraufhin wurde das Reaktionsgemisch *in vacuo* zur

Trockene eingeeengt, der Rückstand wurde ohne weitere Aufreinigung direkt in Methanol (29 mL) gelöst und konzentrierte Schwefelsäure (1.14 mL, 21.4 mmol, 3.6 eq.) wurde langsam zugetropft. Das Gemisch wurde 19 h bei 80 °C verrührt und nach Abkühlen auf Raumtemperatur mit gesättigter wässriger Na₂CO₃-Lösung neutralisiert, filtriert und Lösungsmittel *in vacuo* entfernt. Das Rohprodukt wurde mittels Säulenchromatographie an Silicagel (CH₂Cl₂:MeOH 10:1) aufgereinigt, um 4-Amino-3,5-dihydroxybenzoesäuremethylester **59** (655 mg, 3.58 mmol, 61% über 2 Stufen) als gelben Feststoff zu erhalten.

¹H-NMR (400 MHz, D₃COD): δ/ppm = 7.01 (s, 2H, 2-H), 3.80 (s, 3H, 6-H).

¹³C-NMR (101 MHz, D₃COD): δ/ppm = 169.61 (5-C), 145.77 (3-C), 130.94 (4-C), 119.22 (1-C), 109.62 (2-C), 52.23 (6-C).

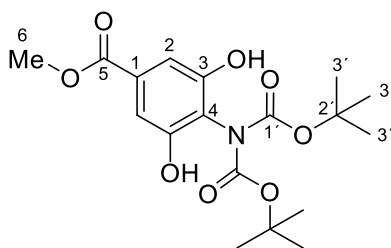
HRMS (ESI⁺): m/z berechnet für C₈H₁₀NO₄⁺ [M+H]⁺ 184.0604; gefunden 184.0604.

HRMS (ESI⁻): m/z berechnet für C₈H₈NO₄⁻ [M-H]⁻ 182.0459; gefunden 182.0459.

IR (ATR): ν (cm⁻¹) = 3402 (w), 3316 (w), 3264 (w), 2951 (w), 2519 (w), 1700 (s), 1612 (m), 1534 (w), 1438 (s), 1351 (m), 1313 (w), 1236 (m), 1222 (m), 1136 (m), 1097 (w), 998 (s), 912 (w), 870 (m), 829 (m), 788 (m), 762 (s), 739 (m).

R_f (CH₂Cl₂:MeOH 10:1) = 0.50.

4-(Di-*tert*-butyloxycarbonyl)amino-3,5-dihydroxybenzoesäuremethylester (**60**)



4-Amino-3,5-dihydroxybenzoesäuremethylester **59** (655 mg, 3.58 mmol, 1.0 eq.) wurde in 23 mL 1,4-Dioxan suspendiert und mit Triethylamin (13.5 mL, 97.4 mmol, 27.2 eq.), Boc₂O (1.80 g, 8.25 mmol, 2.3 eq.) und schließlich mit DMAP (27 mg, 0.22 mmol, 0.1 eq.) versetzt. Das Reaktionsgemisch wurde 15 h bei Raumtemperatur verrührt, anschließend wurde das Gemisch in gesättigte wässrige NaHCO₃-Lösung (150 mL) gegossen und die wässrige Phase wurde mit EtOAc (3 x 150 mL) extrahiert. Die vereinigten organischen Phasen wurden mit gesättigter wässriger

Natriumchloridlösung (100 mL) gewaschen, über Na₂SO₄ getrocknet und *in vacuo* zur Trockene eingengt. Der Rückstand wurde über Säulenchromatographie an Silicagel (iHex/EtOAc 15:1 → 10:1 → 7:1) aufgereinigt, um die an der Aminogruppe zweifach Boc-geschützte Verbindung **60** (1.12 g, 2.92 mmol, 82%) als weißen Schaum zu liefern.

¹H-NMR (400 MHz, CDCl₃): δ/ppm = 7.73 (s, 2H, 2-H), 3.86 (s, 3H, 6-H), 1.56 (s, 18H, 3'-H).

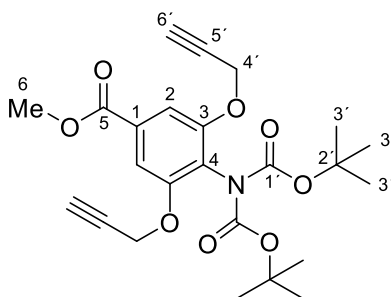
¹³C-NMR (101 MHz, CDCl₃): δ/ppm = 166.12 (5-C), 150.93 (1'-C), 137.99 (3-C), 136.07 (4-C), 121.19 (2-C), 118.43 (1-C), 84.54 (2'-C), 52.22 (6-C), 27.81 (3'-C).

HRMS (ESI⁺): m/z berechnet für C₁₈H₂₆NO₈⁺ [M+H]⁺ 384.1653; m/z gefunden 384.1656.

IR (ATR): ν (cm⁻¹) = 3371 (w), 2982 (w), 1760 (s), 1714 (m), 1626 (s), 1586 (w), 1516 (w), 1438 (m), 1396 (w), 1371 (m), 1338 (m), 1241 (s), 1149 (m), 1113 (s), 1023 (w), 990 (w), 916 (w), 854 (m), 760 (m), 728 (w).

R_f (iHex/EtOAc 3:1) = 0.62.

4-(Di-*tert*-butyloxycarbonyl)amino-3,5-bis(prop-2-yn-1-yloxy)benzoesäuremethylester (**61**)



Verbindung **60** (543 mg, 1.47 mmol, 1.0 eq.) wurde in 15.0 mL trockenem DMF gelöst und mit K₂CO₃ (1.12 g, 8.10 mmol, 5.5 eq.) versetzt. Daraufhin wurde langsam Propargylbromid (870 µL, 8.08 mmol, 5.5 eq.) zugetropft, das Reaktionsgemisch wurde zunächst 8 h bei 80 °C und anschließend 15 h bei Raumtemperatur verrührt. Wegen des unvollständigen Umsatzes des Startmaterials wurde erneut für 4.5 h auf 80 °C erhitzt. Nach Abkühlen auf Raumtemperatur wurde das Gemisch in Wasser (100 mL) gegossen und mit CH₂Cl₂ extrahiert (200 mL). Die organische Phase wurde

mit Wasser gewaschen (3 x 100 mL) und über Na₂SO₄ getrocknet. Lösungsmittel wurden *in vacuo* entfernt und das Rohprodukt säulenchromatographisch an Silicagel (iHex/EtOAc 12:1 → 10:1 → 8:1) aufgereinigt, um die doppelt alkylierte Verbindung **61** (329 mg, 0.716 mmol, 49%) in Form eines viskosen, gelben Öls zu liefern.

¹H-NMR (600 MHz, CDCl₃): δ/ppm = 7.41 (s, 2H, 2-H), 4.76 (d, *J* = 2.4 Hz, 4H, 4'-H), 3.94 (s, 3H, 6-H), 2.50 (t, *J* = 2.4 Hz, 2H, 6'-H), 1.38 (s, 18H, 3'-H).

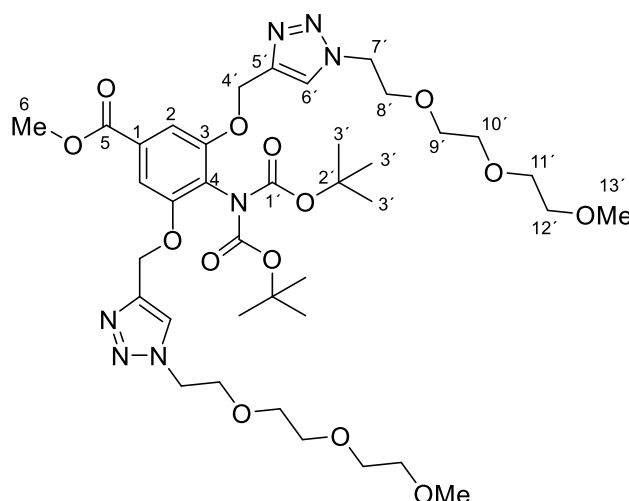
¹³C-NMR (151 MHz, CDCl₃): δ/ppm = 166.51 (5-C), 153.41 (3-C), 150.73 (1'-C), 130.02 (4-C), 123.12 (1-C), 107.44 (2-C), 82.49 (2'-C), 78.00 (5'-C), 76.25 (6'-C), 56.55 (4'-C), 52.63 (6-C), 27.93 (3'-C).

HRMS (ESI⁺): *m/z* berechnet für C₁₆H₁₄NO₈⁺ [M-2C₄H₈+H]⁺ 348.0714; *m/z* gefunden 348.0714.

IR (ATR): ν (cm⁻¹) = 3261 (w), 2981 (w), 2123 (w), 1793 (m), 1752 (w), 1720 (s), 1592 (w), 1436 (m), 1369 (m), 1330 (m), 1278 (m), 1249 (s), 1155 (s), 1107 (s), 1005 (w), 853 (w), 769 (w).

R_f (iHex/EtOAc 3:1) = 0.30.

4-(Di-*tert*-butyloxycarbonyl)amino-3,5-bis((1-(2-(2-(2-methoxyethoxy)ethoxy)ethyl)-1*H*-1,2,3-triazol-4-yl)methoxy)benzoessäuremethylester (62**)**



Ein Gemisch aus THF (5.2 mL) und ddH₂O (5.2 mL) wurde gründlich entgast (*freeze-pump-thaw* 3x) und mit Verbindung **61** (311 mg, 0.677 mmol, 1.0 eq.) und 1-Azido-2-(2-(2-methoxyethoxy)ethoxy)ethan (384 mg, 2.03 mmol, 3.0 eq.) versetzt. Anschließend wurde CuBr·Me₂S (42 mg, 0.20 mmol, 0.3 eq.) zugegeben und das

Reaktionsgemisch wurde 16 h bei Raumtemperatur verrührt. Lösungsmittel wurden *in vacuo* entfernt und der Rückstand über Säulenchromatographie an Silicagel (CH₂Cl₂:MeOH 50:1 → 40:1) aufgereinigt, um das Cycloadditionsprodukt **62** (475 mg, 0.567 mmol, 84%) in Form eines gelben Wachses zu liefern.

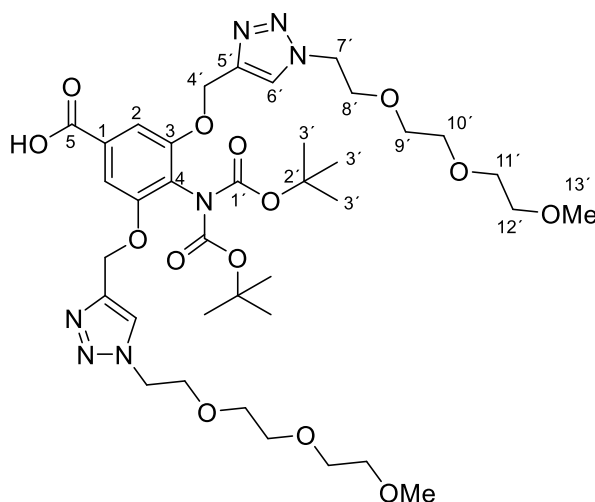
¹H-NMR (600 MHz, CDCl₃): δ/ppm = 7.77 (s, 2H, 6'-H), 7.43 (s, 2H, 2-H), 5.28 (s, 4H, 4'-H), 4.54 (t, *J* = 5.2 Hz, 4H, 7'-H), 3.93 (s, 3H, 6-H), 3.88 (t, *J* = 5.2 Hz, 4H, 8'-H), 3.63 – 3.50 (m, 16H, 9'-H, 10'-H, 11'-H, 12'-H), 3.36 (s, 6H, 13'-H), 1.30 (s, 18H, 3'-H).

¹³C-NMR (151 MHz, CDCl₃): δ/ppm = 123.84 (6'-C), 106.96 (2-C), 72.04 – 70.39 (9'-C, 10'-C, 11'-C, 12'-C), 69.38 (8'-C), 63.54 (4'-C), 59.03 (13'-C), 52.54 (6-C), 50.46 (7'-C), 27.80 (3'-C)

HRMS (ESI⁺): *m/z* berechnet für C₃₈H₆₀N₇O₁₄⁺ [M+H]⁺ 838.4193; *m/z* gefunden 838.4185.

R_f (CH₂Cl₂:MeOH 20:1) = 0.42.

4-(Di-*tert*-butyloxycarbonyl)amino-3,5-bis((1-(2-(2-(2-methoxyethoxy)ethoxy)-ethyl)-1*H*-1,2,3-triazol-4-yl)methoxy)benzoesäure (63)



Methylester **62** (254 mg, 0.303 mmol, 1.0 eq.) wurde in 3.0 mL MeOH gelöst und eine wässrige Lösung von LiOH·H₂O (64 mg, 1.52 mmol, 5.0 eq.) in ddH₂O (0.65 mL) wurde zugetropft. Es wurde 17 h bei Raumtemperatur verrührt, anschließend wurde Methanol *in vacuo* entfernt, der Rückstand wurde in gesättigter wässriger NH₄Cl-Lösung (100 mL) aufgenommen und mit CH₂Cl₂ extrahiert (3 x 100 mL). Die

wässrige Phase wurde daraufhin mit wässriger Salzsäure (6 M) auf pH 3 angesäuert und nochmals mit CH₂Cl₂ (100 mL) extrahiert. Die vereinigten organischen Phasen wurden mit einer wässrigen 1 M EDTA-Lösung (50 mL) und mit gesättigter wässriger Natriumchloridlösung (50 mL) gewaschen und über Na₂SO₄ getrocknet. Lösungsmittel wurden *in vacuo* entfernt und der Rückstand unter Hochvakuum getrocknet, um Verbindung **63** (220 mg, 0.267 mmol, 88%) als gelbes Wachs zu erhalten, das ohne weitere Aufreinigung für die nächste Stufe verwendet werden konnte.

¹H-NMR (600 MHz, CDCl₃): δ/ppm = 7.78 (s, 2H, 6'-H), 7.47 (s, 2H, 2-H), 5.33 (s, 4H, 4'-H), 4.53 (t, *J* = 5.1 Hz, 4H, 7'-H), 3.86 (t, *J* = 5.2 Hz, 4H, 8'-H), 3.63–3.54 (m, 16H, 9'-H, 10'-H, 11'-H, 12'-H), 3.39 (s, 6H, 13'-H), 1.31 (s, 18H, 3'-H).

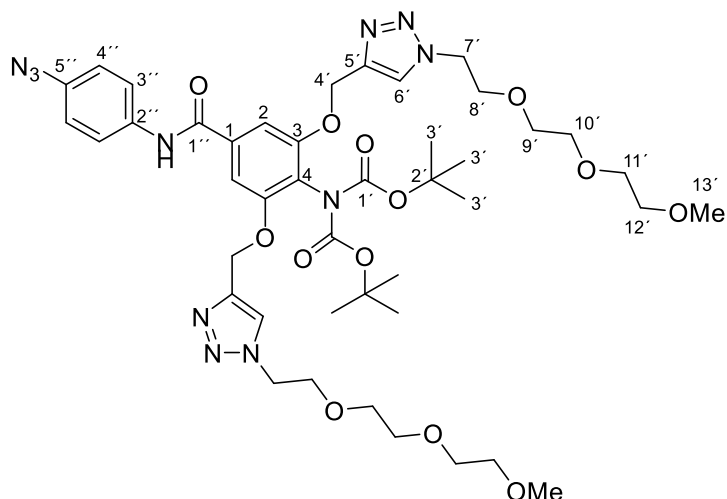
¹³C-NMR (151 MHz, CDCl₃): δ/ppm = 167.84 (5-C), 154.09 (3-C), 151.06 (1'-C), 143.51 (5'-C), 130.39 (4-C), 124.12 (6'-C), 122.97 (1-C), 107.68 (2-C), 82.50 (2'-C), 72.08 – 70.56 (9'-C, 10'-C, 11'-C, 12'-C), 69.60 (8'-C), 63.16 (4'-C), 59.12 (13'-C), 50.59 (7'-C), 27.93 (3'-H).

HRMS (ESI⁺): *m/z* berechnet für C₃₇H₅₇N₇NaO₁₄⁺ [M+Na]⁺ 846.3856; *m/z* gefunden 846.3853.

HRMS (ESI⁻): *m/z* berechnet für C₃₇H₅₆N₇O₁₄⁻ [M-H]⁻ 822.3891; *m/z* gefunden 822.3891.

R_f (CH₂Cl₂:MeOH 20:1) = 0.38.

***N*-(4-Azidophenyl)-4-(di-*tert*-butyloxycarbonyl)amino-3,5-bis((1-(2-(2-(2-methoxyethoxy)ethoxy)ethyl)-1*H*-1,2,3-triazol-4-yl)methoxy)benzamid (**64**)**



Carbonsäure **63** (183 mg, 0.22 mmol, 1.0 eq.) wurde in 3.0 mL trockenem DMF gelöst und mit 4-Azidoanilin-Hydrochlorid (38 mg, 0.22 mmol, 1.0 eq.) versetzt. Im Anschluss wurden DIPEA (60 μ L, 0.29 mmol, 1.3 eq.) und HBTU (110 mg, 0.29 mmol, 1.3 eq.) hinzugefügt und das Reaktionsgemisch wurde 16 h bei Raumtemperatur verrührt. Das Lösungsmittel wurde *in vacuo* entfernt und das Rohprodukt mittels Säulenchromatographie an Silicagel ($\text{CH}_2\text{Cl}_2 \rightarrow \text{CH}_2\text{Cl}_2:\text{MeOH}$ 40:1 $\rightarrow \text{CH}_2\text{Cl}_2:\text{MeOH}$ 30:1) aufgereinigt, um das Kupplungsprodukt **64** (113 mg, 0.12 mmol, 55%) als braunes Wachs zu liefern.

$^1\text{H-NMR}$ (600 MHz, CDCl_3): δ/ppm = 8.43 (s, 1H, CO-NH), 7.79 (s, 2H, 6'-H), 7.76 (d, J = 8.8 Hz, 2 H, 3''-H), 7.38 (s, 2H, 2-H), 7.04 (d, J = 8.9 Hz, 2H, 4''-H), 5.31 (s, 4H, 4'-H), 4.52 (t, J = 5.1 Hz, 4H, 7'-H), 3.86 (t, J = 5.2 Hz, 4H, 8'-H)), 3.61–3.51 (m, 16H, 9'-H, 10'-H, 11'-H, 12'-H), 3.36 (s, 6H, 13'-H), 1.33 (s, 18 H, 3'-H).

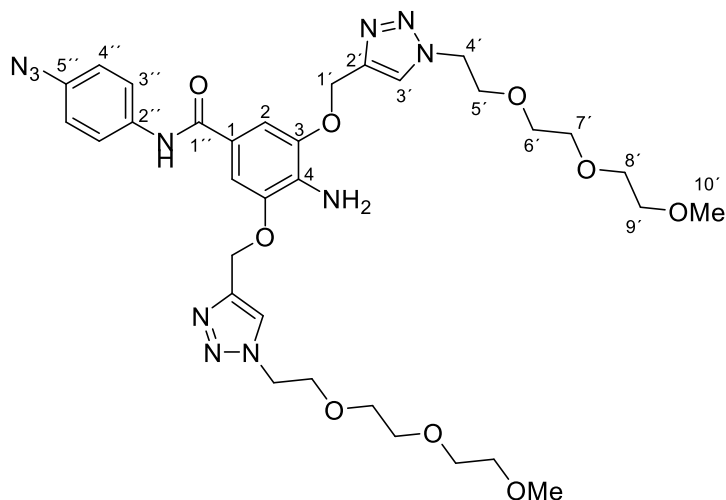
$^{13}\text{C-NMR}$ (151 MHz, CDCl_3): δ/ppm = 165.08 (1''-C), 154.26 (3-C), 151.28 (1'-C), 143.43 (5'-C), 135.93 (5''-C), 135.51 (2''-C), 135.07 (4-C), 124.50 (6'-C), 121.75 (7-C), 121.65 (1-C), 119.67 (8-C), 105.64 (2-C), 82.08 (2'-C), 72.03 – 69.48 (9'-C, 10'-C, 11'-C, 12'-C), 62.91 (4'-C), 59.17 (13'-C), 50.51 (7'-C), 27.99 (3'-C).

HRMS (ESI⁺): m/z berechnet für $\text{C}_{43}\text{H}_{62}\text{N}_{11}\text{O}_{13}^+$ $[\text{M}+\text{H}]^+$ 940.4523; m/z gefunden 940.4517.

HRMS (ESI⁻): m/z berechnet für $\text{C}_{44}\text{H}_{62}\text{N}_{11}\text{O}_{15}^-$ $[\text{M}+\text{HCOO}]^-$ 984.4432; m/z gefunden 984.4430.

R_f ($\text{CH}_2\text{Cl}_2:\text{MeOH}$ 20:1) = 0.33.

***N*-(4-Azidophenyl)-4-amino-3,5-bis((1-(2-(2-(2-methoxyethoxy)ethoxy)ethyl)-1*H*-1,2,3-triazol-4-yl)methoxy)benzamid (65)**



Die zweifach Boc-geschützte Verbindung **64** (187 mg, 0.199 mmol, 1.0 eq.) wurde in 5.0 mL trockenem CH₂Cl₂ gelöst und bei 0 °C tropfenweise mit Trifluoressigsäure (0.60 mL, 7.8 mmol, 39 eq.) versetzt. Das Reaktionsgemisch wurde 4 h bei Raumtemperatur verrührt, anschließend wurde die Reaktion durch Zugabe von Wasser (30 mL) und Einstellen des pH auf 8 unter Verwendung von wässriger NaOH (2 M) gestoppt. Daraufhin wurde mit CH₂Cl₂ extrahiert (3 x 50 mL), die vereinigten organischen Phasen wurden mit gesättigter wässriger Natriumchloridlösung (30 mL) gewaschen und über Na₂SO₄ getrocknet. Das Lösungsmittel wurde *in vacuo* entfernt und das Rohprodukt mittels Säulenchromatographie an Silicagel (CH₂Cl₂:MeOH 40:1) aufgereinigt, um die entschützte Verbindung **65** (100 mg, 0.135 mmol, 68%) in Form eines braunen Wachses zu erhalten.

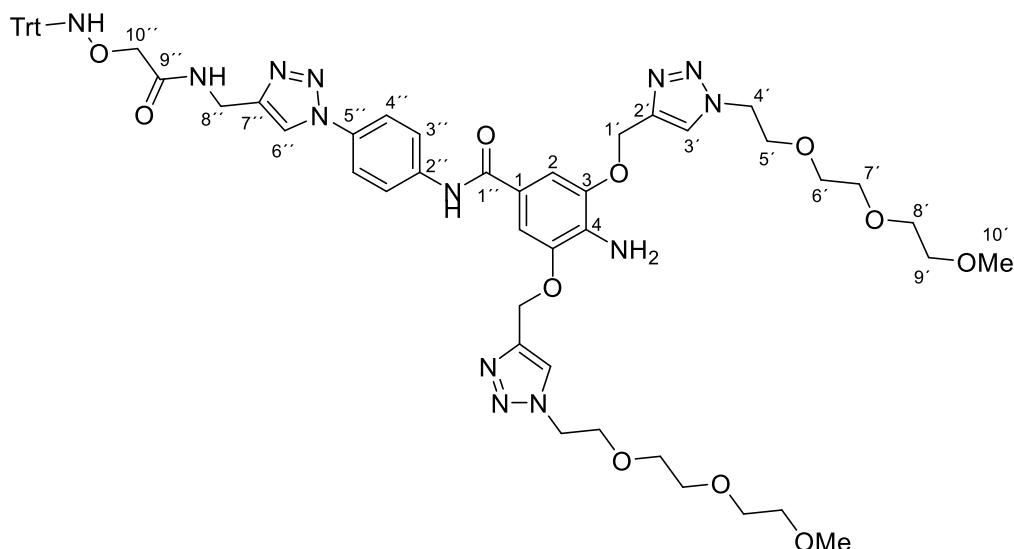
¹H-NMR (600 MHz, CDCl₃): δ/ppm = 8.19 (s, 1H, CO-NH), 7.87 (s, 2H, 3'-H), 7.72 (d, *J* = 8.9 Hz, 2H, 3''-H), 7.36 (s, 1.9H, 2-H), 7.02 (d, *J* = 8.9 Hz, 2H, 4''-H), 5.30 (s, 2H, NH₂), 5.27 (s, 4H, 1'-H), 4.55 (t, *J* = 5.5 Hz, 4 H, 4'-H), 3.86 (t, *J* = 5.5 Hz, 4H, 5'-H), 3.61 – 3.51 (m, 16H, 6'-H, 7'-H, 8'-H, 9'-H), 3.34 (s, 6H, 10'-H).

¹³C-NMR (101 MHz, CDCl₃): δ/ppm = 165.72 (1''-C), 145.11 (3-C), 143.55 (2'-C), 135.97 (2''-C), 135.42 (5''-C), 131.32 (4-C), 124.72 (3'-C), 122.19 (1-C), 121.68 (3''-C), 119.62 (4''-C), 106.53 (2-C), 72.05 – 70.60 (6'-C, 7'-C, 8'-C, 9'-C), 69.48 (5'-C), 62.65 (1'-C), 59.13 (10'-C), 50.50 (4'-C).

HRMS (ESI⁺): *m/z* berechnet für C₃₃H₄₆N₁₁O₉⁺ [M+H]⁺ 740.3475; *m/z* gefunden 740.3477.

R_f (CH₂Cl₂:MeOH 10:1) = 0.65.

4-Amino-3,5-bis((1-(2-(2-(2-methoxyethoxy)ethoxy)ethyl)-1H-1,2,3-triazol-4-yl)-methoxy)-N-(4-(4-((2-((tritylamino)oxy)acetamido)methyl)-1H-1,2,3-triazol-1-yl)-phenyl)benzamid (66)



Ein Gemisch aus THF (2.0 mL) und ddH₂O (2.0 mL) wurde gründlich entgast (*freeze-pump-thaw* 3x) und mit Azid **66** (87 mg, 118 μ mol, 1.0 eq.) und Alkin **41** (46 mg, 129 μ mol, 1.1 eq.) versetzt. Daraufhin wurde CuBr·Me₂S (11 mg, 54 μ mol, 0.4 eq.) hinzugefügt und das Reaktionsgemisch 16 h bei Raumtemperatur verrührt. Anschließend wurde das Gemisch *in vacuo* zur Trockene eingeeengt und der Rückstand mittels Säulenchromatographie an Silicagel (CH₂Cl:MeOH 50:1 → CH₂Cl:MeOH 40:1 → CH₂Cl:MeOH 30:1) aufgereinigt, um das Cycloadditionsprodukt **66** (100 mg, 90 μ mol, 76%) in Form eines braunen Wachses zu erhalten.

¹H-NMR (400 MHz, CDCl₃): δ /ppm = 8.45 (s, 1H, CO-NH-Ar), 7.92 – 7.88 (m, 4H, 3'-H, 3''-H), 7.80 (s, 1H, 6''-H), 7.66 (d, J = 8.9 Hz, 2H, 4''-H), 7.41 (s, 2H, 2-H), 7.26 (s, 15H, CPh₃), 6.35 (t, J = 5.9 Hz, 1H, CO-NH-Alk), 5.28 (s, 4H, 1'-H), 4.55 (t, J = 4.9 Hz, 4H, 4'-H), 4.44 (d, J = 5.9 Hz, 2H, 8''-H), 4.28 (s, 2H, 10''-H), 3.87 (t, J = 5.0 Hz, 4H, 5'-H), 3.63 – 3.49 (m, 16H, 6'-H, 7'-H, 8'-H, 9'-H), 3.34 (s, 6H, 10'-H)

¹³C-NMR (101 MHz, CDCl₃): δ /ppm = 169.67 (9''-C), 165.88 (1''-C), 145.33 (3-C), 145.26 (7''-C), 143.81 (CPh₃), 143.44 (2'-C), 139.51 (2''-C), 132.71 (5''-C), 130.66 (4-C), 129.01 (CPh₃), 128.12 (CPh₃), 128.07 (CPh₃), 127.34 (CPh₃), 124.84 (1-C), 121.29 (4''-C), 120.95 (3''-C), 120.49 (6''-C), 106.65 (2-C), 74.53 (CPh₃), 73.61

(10''-C), 72.05 – 70.60 (6'-C, 7'-C, 8'-C, 9'-C), 69.46 (5'-C), 62.65 (1'-C), 59.14 (10'-C), 50.54 (4'-C), 34.64 (8''-C).

HRMS (ESI⁺): m/z berechnet für C₅₇H₆₈N₁₃O₁₁⁺ [M+H]⁺ 1110.5156; m/z gefunden 1110.5156.

IR (ATR): ν (cm⁻¹) = 3304 (w), 2920 (w), 1648 (m), 1593 (w), 1520 (s), 1442 (m), 1408 (w), 1330 (m), 1260 (m), 1196 (w), 1143 (s), 1099 (s), 1046 (s), 804 (s), 753 (m), 699 (s).

R_f (CH₂Cl₂:MeOH 20:1) = 0.50.

Bindungsstudien der Hydroxylaminreagenzien mit fdC auf Oligonukleotidebene

Die folgende Tabelle zeigt einen Ansatz einer Bindungsstudie, die die Reaktion zu sechs Zeitpunkten untersuchte. Das DNA-Oligonukleotid (5'-UUU UG^fC GGU UG-3') wurde dem Reaktionsgemisch als letztes zugefügt, anschließend wurde das Reaktionsgemisch bei 37 °C inkubiert und Proben nach definierten Zeitpunkten entnommen (hier: 5 min, 2 h, 4 h, 6 h, 24 h).

Tabelle 6-1: Reaktionsbedingungen zur Derivatisierung von fdC-DNA mit den synthetisierten Hydroxylaminreagenzien.

fdC-Bindungsstudie	konz. Stocklsg.	Reaktionskonz.	V [μL]	V [μL]
			1x	6x
Hydroxylaminreagenz	10.0 mM	120 nmol	12	72
HEPES	150 mM		3.99	23.94
EDTA	1 mM		3	18
fdC-DNA	68.6 μM	400 pmol	5.83	34.98
Anisidin	35.73 mM	8 mM = 160 nmol	4.48	26.88
dd-Wasser			0.7	4.2
Gesamtvolumen			30	180

Nach der Probenentnahme (jeweils 30 μL) wurde die Reaktion durch Zugabe von 5 μL Aceton (HPLC-Güte) gestoppt und dem Gemisch wurden 20 μL einer wässrigen Lösung von HNEt₃OAc (0.1 M, pH = 7) zugefügt. Das Gemisch wurde sofort mittels analytischer HPLC untersucht (0% Puffer B auf 30% Puffer B in 45 min).

Puffer A: 0.1 M HNEt₃OAc in ddH₂O (pH = 7)

Puffer B: 0.1 M HNEt₃OAc in 80% aq. MeCN (pH = 7)

6.5. Synthese eines carbazyklischen Derivats von 5-Azacytidin und Untersuchungen zu dessen Einfluss auf *in vivo* Methylierung von DNA

6.5.1. Synthese eines carbazyklischen 5-Azacytidins

Kürzlich konnte von *Wildenhof et al.*^[32] gezeigt werden, dass ein carbazyklisches Derivat von 5-Aza-2'-desoxycytidin (cAza-dC) im Gegensatz zur 2'-Desoxyribose-basierten Verbindung 5-Aza-2'-desoxycytidin (5-Aza-dC, Decitabin) keine Hydrolyseanfälligkeit in wässriger Umgebung bei physiologischem pH zeigt. Gleichzeitig weist es eine nur leicht verminderte Aktivität zur Inhibierung der DNA-Methylierung auf. Auf Basis dieser Erkenntnisse rückte natürlich die Frage in den Vordergrund der Forschung, ob ein carbazyklisches Derivat des 5-Azacytidins (cAzaC) auch solche Eigenschaften besitzen könnte. Die Ribose-basierte, kommerziell erhältliche Verbindung 5-Azacytidin (5-Aza-C) wird klinisch unter dem Namen Azacitidin zur Therapie von AML und MDS eingesetzt und weist die für 5-Aza-Cytidine charakteristische Hydrolyseanfälligkeit in wässriger Umgebung auf.^[33] Ein möglicherweise hydrolysestabiles Derivat dieser Verbindung, welches gleichzeitig seine Aktivität zur Inhibierung der *in vivo* DNA-Methylierung beibehalten würde, wirft großes Interesse im Hinblick auf seine medizinisch-chemische Erforschung auf.

Die in Abbildung 6-31 dargestellte Synthese von cAzaC erfolgte in Anlehnung an die Arbeiten von *Wildenhof et al.*,^[32] wobei hier allerdings zunächst ausgehend vom Boc-geschützten (1*R*)-*Vince*-Lactam **70** eine Bishydroxylierung durchgeführt wurde. Um bei der Zielverbindung der Synthese die gewünschte Stereochemie zu erreichen, d.h. um eine abschließende *ribo*-Konfiguration der Substituenten am Cyclopentangrundgerüst und keine *lyxo*-Konfiguration im Zielmolekül zu erhalten, war hier eine selektive *exo*-Bishydroxylierung des Norbornen-Grundgerüsts notwendig. Die *exo*-Selektivität der Bishydroxylierungsreaktion am *Vince*-Lactam-Grundgerüst ist bereits in der Literatur beschrieben.^[132] Die Boc-Schutzgruppe an der Amidfunktion sollte dabei zusätzlich die *exo*-Selektivität noch weiter steigern, da *Wildenhof et al.* zeigen konnten, dass zumindest bei einer Epoxidierung desselben Grundgerüsts die Boc-Schutzgruppe an jener Position die *exo*-Selektivität der Epoxidierung verbessern konnte.^[32]

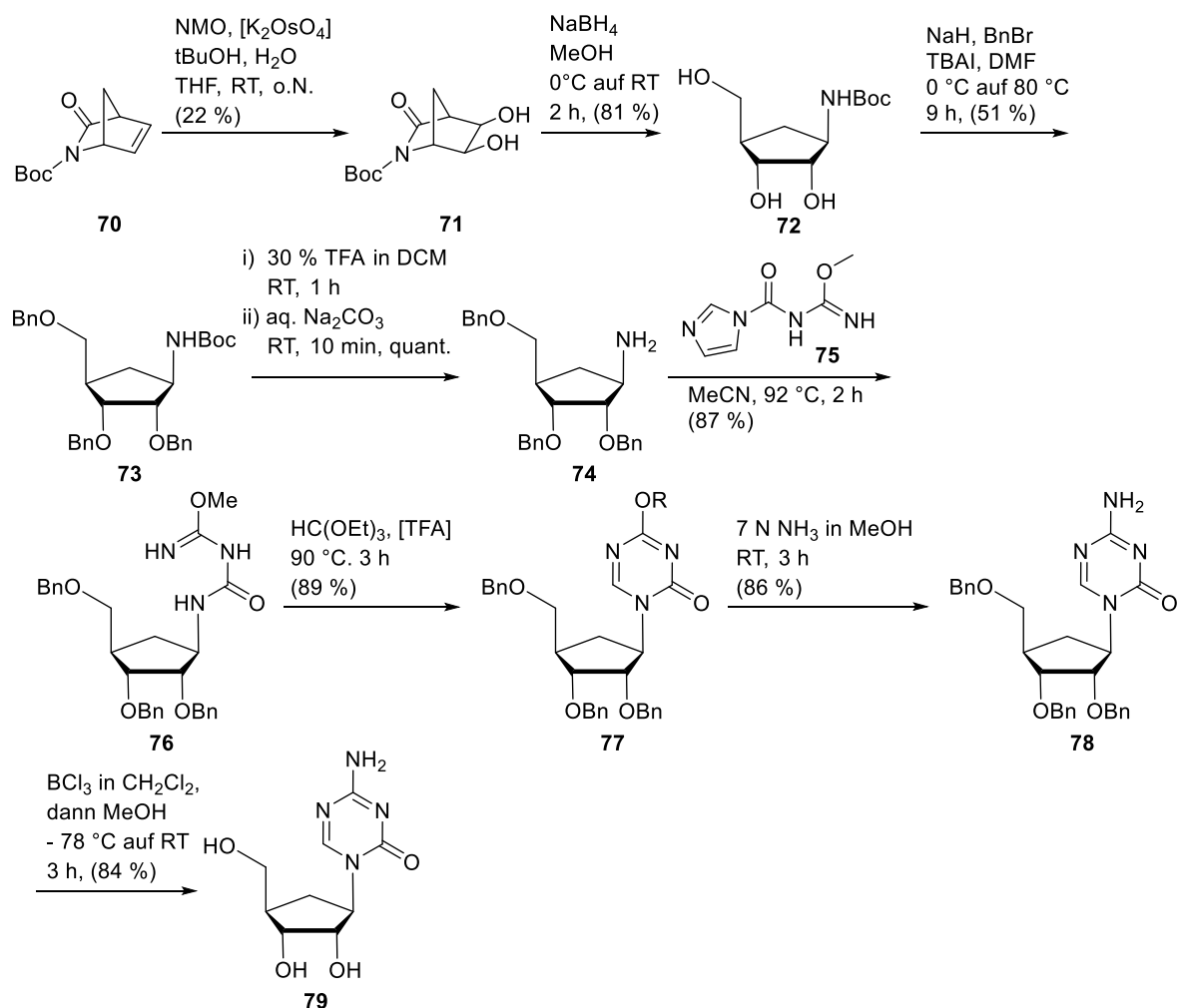


Abbildung 6-31: Schema der Synthese von cAzaC **79**, R = Me, Et.

Zur Bishydroxylierung wurde die *Upjohn*-Variante der Osmiumtetroxid-Bishydroxylierung von Alkenen durchgeführt, um große Mengen des äußerst giftigen Osmiumtetroxids zu vermeiden.^[104] Allerdings konnte in diesem Fall das gewünschte *exo*-Bishydroxylierungsprodukt in nur 22% Ausbeute isoliert werden, ein *endo*-Bishydroxylierungsprodukt konnte nicht detektiert und isoliert werden.

Zur Bestätigung der bis-*exo*-Konfiguration der neu eingeführten Hydroxygruppen wurde eine genaue NMR-spektroskopische Analyse an Hand der Kopplungskonstanten der Signale des ^1H -NMR-Spektrums durchgeführt. Abbildung 6-32 zeigt den ausgewählten Bereich des NMR-Spektrums der Verbindung **71**.

Kam et al. berichteten, dass bei einer *exo*-Bishydroxylierung des nicht-Boc-geschützten *Vince*-Lactams die Kopplungskonstante der vicinalen 5-6-Kopplung im ^1H -NMR-Spektrum im Bereich zwischen 5-6 Hz und die Kopplungskonstanten der vicinalen 1-6- und 4-5-Kopplung im Bereich <2 Hz liegen, wohingegen bei einer

endo-Bishydroxylierung alle beobachteten Kopplungskonstanten auf Grund der sehr kleinen Diederwinkel zwischen den Protonen der 1-, 4-, 5-, und 6-Positionen im Bereich von 5-6 Hz lägen.^[132] Wie in Abbildung 6-32 ersichtlich können bei den Signalen der Protonen an 5- und 6-Position Kopplungskonstanten von ca. 6 Hz detektiert werden, wohingegen bei der 1- und 4-Position nur äußerst kleine Kopplungskonstanten beobachtet werden. Dies bestätigt die erfolgreiche Synthese des *exo*-Bishydroxylierungsproduktes **71**.

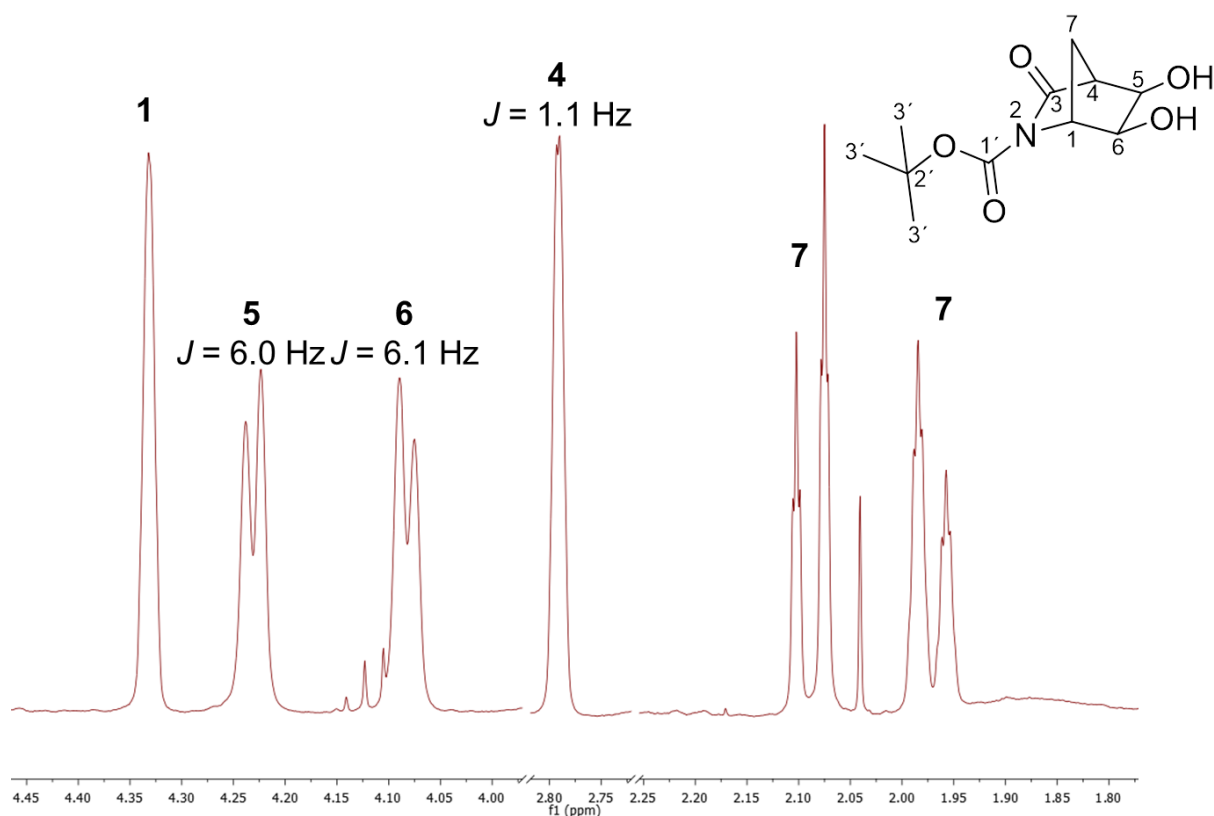


Abbildung 6-32: ¹H-NMR Spektrum eines ausgewählten Bereichs des Bishydroxylierungsproduktes **71** einschließlich der relevanten Kopplungskonstanten.

Nach erfolgter Bishydroxylierung wurde anschließend die Amidbindung der Verbindung **71** reaktiv zum Aminoalkohol unter Verwendung von Natriumborhydrid gespalten. Aufgrund der stereochemischen Eigenschaften der Ausgangsverbindung konnte somit ein Aminoalkohol-substituiertes Cyclopentan **72** erhalten werden, welches eine *ribo*-Konfiguration nachstellte. Im Anschluss wurden alle Hydroxygruppen der Verbindung unter Verwendung von Benzylbromid in basischer Umgebung Benzyl-geschützt und daraufhin wurde die Boc-geschützte Aminofunktion unter stark sauren Bedingungen entschützt, um nach Neutralisation ein ungeschütztes Aminoalkohol-substituiertes Cyclopentanderivat **74** zu erhalten. Dieses wurde auf Grund seiner Oxidationsempfindlichkeit nach Trocknung unter

Hochvakuum ohne weitere Aufreinigung direkt mit dem Imidazol-aktivierten Isoharnstoffderivat **75** umgesetzt.^[32] Das dabei erhaltene Methylisobiuretderivat **76** wurde im weiteren Verlauf mit Triethylorthoformiat unter Säurekatalyse umgesetzt, um eine Zyklisierung zum 1,3,5-Triazinring in Form eines trennbaren Gemisches aus 4-Methoxy- und 4-Ethoxytriazin **77** zu erreichen. Nachfolgende Aminierung der beiden Verbindungen in separaten Ansätzen oder in Form eines Gemisches der Verbindungen unter Verwendung von methanolischem Ammoniak lieferte das geschützte, carbazyklische 5-Azacytidinderivat **78**, welches abschließend unter Verwendung von Bortrichlorid in Dichlormethan entschützt wurde. Die durchgeführten synthetischen Transformationen resultierten nach der Bishydroxylierung durchgehend in annehmbaren bis guten Ausbeuten, wodurch mehrere hundert Milligramm des carbazyklischen 5-Azacytidins **79** erhalten werden konnten.

Eine Umkristallisation aus Methanol zur Aufreinigung des Rohprodukts analog zu den Arbeiten von *Wildenhof et al.* konnte hier allerdings nicht erfolgreich durchgeführt werden,^[32] da die erhaltene Verbindung **79** eine sehr gute Löslichkeit in Methanol aufwies. Umkristallisationen aus Ethanol, Isopropanol oder Gemischen dieser schlugen auf Grund der vollständigen Unlöslichkeit der Verbindung **79**, auch in großen Mengen der untersuchten Lösungsmittel und unter langem Erwärmen, fehl. Aus diesem Grund konnten keine Kristalle der Verbindung erhalten und keine Strukturaufklärung mittels Röntgenstrukturanalyse durchgeführt werden.

Eine Aufreinigung mittels HPL-Chromatographie zur Aufreinigung der Verbindung **79** für geplante zellbiologische Verfütterungsexperimente erschien möglich, allerdings musste hierfür zunächst die Stabilität der Verbindung in wässriger Umgebung gezeigt werden.

6.5.2. Untersuchungen zur Hydrolysestabilität der carbazyklischen Verbindung

Um die Hydrolyseeigenschaften der carbazyklischen Verbindung **79** im Vergleich zum bekannten 5-Aza-C zu untersuchen, wurden beide Verbindungen bei Raumtemperatur in wässriger Umgebung für einen Zeitraum von mehreren Tagen inkubiert und zu ausgewählten Zeitpunkten mittels analytischer HPL-Chromatographie untersucht. Während bei der Ribose-basierten Verbindung 5-Aza-C schon innerhalb kurzer Zeit verschiedene Hydrolyseprodukte nachgewiesen werden konnten (Abbildung 6-33 A), erwies sich die carbazyklische Verbindung als hydrolysestabil und es konnte auch nach mehreren Tagen Inkubation in wässriger

Lösung ausschließlich das Startmaterial cAzaC detektiert werden (Abbildung 6-33 B). Leichte Verschiebungen der Retentionszeiten der untersuchten Verbindungen in den dargestellten Chromatogrammen können auf die Verwendung leicht unterschiedlich konditionierter Chromatographiesäulen zu den verschiedenen Zeitpunkten zurückgeführt werden.

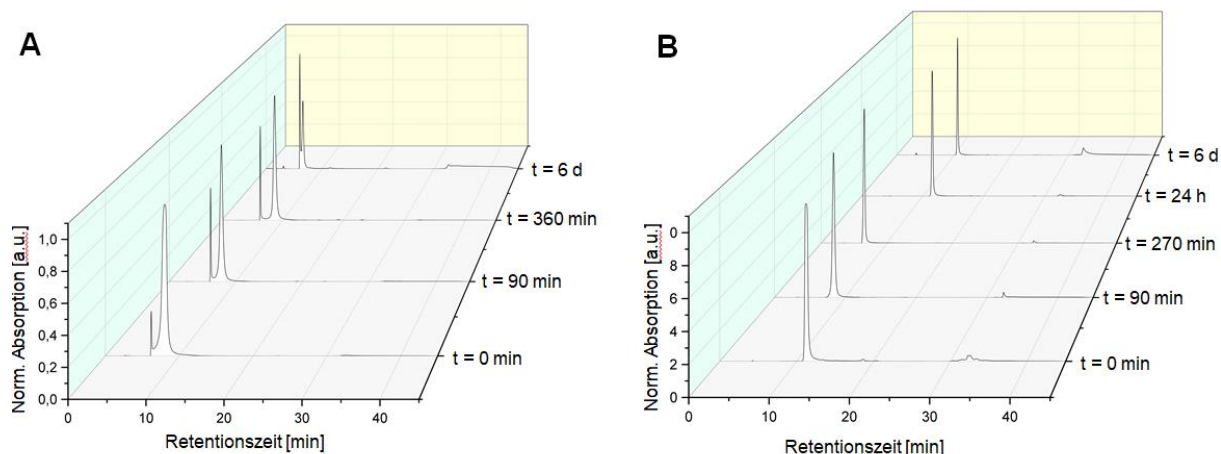


Abbildung 6-33: Ergebnisse der Hydrolysestabilitätsuntersuchungen von A) 5-Aza-C und B) cAzaC.

Auf Basis der erhaltenen Ergebnisse wurde eine semipräparative HPL-chromatographische Aufreinigung der synthetisierten Verbindung **79** durchgeführt, um eine ausreichende Reinheit der Verbindung für zellbiologische Verfütterungsexperimente zu erreichen.

Des Weiteren konnte mit diesen Experimenten gezeigt werden, dass bei diesen Zellverfütterungsexperimenten von einer Hydrolysestabilität der eingesetzten Verbindung unter physiologischem pH ausgegangen werden kann. Auf Basis der Hydrolysestabilität der synthetisierten carbazyklischen Verbindung **79** sollten nun Verfütterungsexperimente an Zellkulturen durchgeführt werden, um den Einfluss von cAzaC auf die *in vivo*-Methylierung zu untersuchen.

6.5.3. Zellverfütterungsexperimente mit c-AzaC und erhaltene Ergebnisse

Da bekannt ist, dass Ribonukleotide auf Di- und Triphosphatebene *in vivo* enzymatisch in die korrespondierenden 2'-Desoxyribonukleotide umgewandelt werden können,^[133] wurde im Rahmen der hier durchgeführten Zellverfütterungsexperimente zunächst der Einfluss der synthetisierten Verbindung cAzaC auf die DNA-Methylierung untersucht. Dies schien auf Grund des erwarteten Effekts der Verbindung auf die DNA-Methylierung sinnvoll.

Die zellbiologischen Verfütterungsexperimente wurden von *Dr. Franziska Traube* und die anschließenden massenspektrometrischen Messungen und zugehörigen Datenauswertungen wurden von *Florian Schelter* in Zusammenarbeit mit *Dr. Franziska Traube* durchgeführt, weswegen hier nur ein kurzer Überblick über die durchgeführten Experimente und die erhaltenen Ergebnisse gegeben werden soll.

Im Rahmen der Zellverfütterungsexperimente wurde das neu synthetisierte cAzaC, das nach *Wildenhof et al.* synthetisierte cAza-dC,^[32] sowie die kommerziell erhältlichen Verbindungen 5-Aza-C und 5-Aza-dC in verschiedenen Konzentrationen an die Leukämie-Zelllinien AML-491 und MOLM-13 über mehrere Tage verfüttert, anschließend wurde die DNA der Zellen isoliert und verdaut. Das erhaltene Nukleosidgemisch wurde massenspektrometrisch im Hinblick auf den Gehalt an mdC und hmdC analysiert.

Abbildung 6-34 zeigt exemplarisch, dass bei der Verfütterung von cAzaC an MOLM-13 Zellen in verschiedenen Konzentrationen kein Effekt auf den detektierten Gehalt von mdC im Vergleich zur ungefütterten Kontrolle festgestellt werden konnte. Verfütterungsexperimente von cAza-dC, 5-Aza-dC und 5-Aza-C hingegen resultierten allesamt in einer verringerten Abundanz von mdC in der DNA der untersuchten Zelllinie im Vergleich zum ungefütterten Kontrollexperiment.

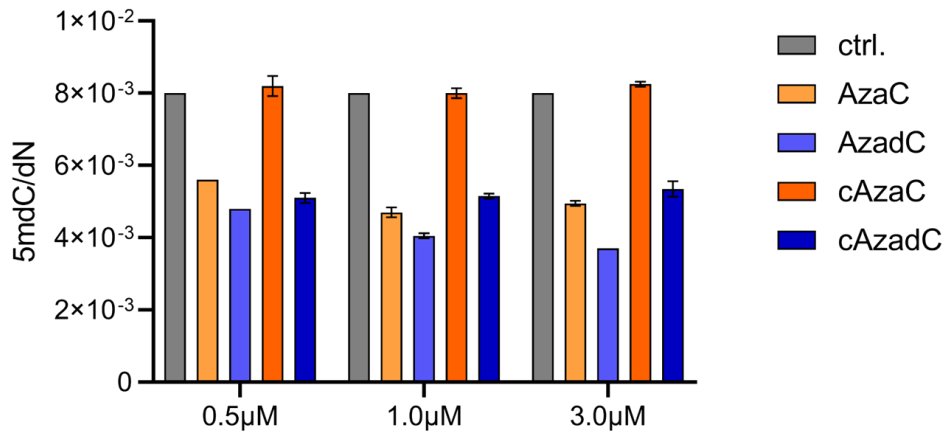


Abbildung 6-34: Ergebnisse der massenspektrometrischen Quantifizierung von mdC bei Zellverfütterung von AzaC, AzadC, cAzaC und cAzadC im Vergleich zur ungefütterten Kontrolle.

Ein ähnliches Resultat konnte bei Untersuchungen im Hinblick auf den hmdC-Gehalt in der DNA bei Zellverfütterungsexperimenten nachgewiesen werden. Auch bei der Zellverfütterung von cAzaC in verschiedenen Konzentrationen an Zellen der AML-491 Zelllinie konnte ebenfalls kein reduzierter Gehalt an mdC in den untersuchten DNA-Proben im Vergleich zur ungefütterten Kontrollprobe zur Folge, wohingegen Verfütterungsexperimente mit cAza-dC, 5-Aza-dC oder 5-Aza-C in einer deutlich verringerten Abundanz dieser DNA-Modifikation resultierten (siehe Abbildung 6-35).

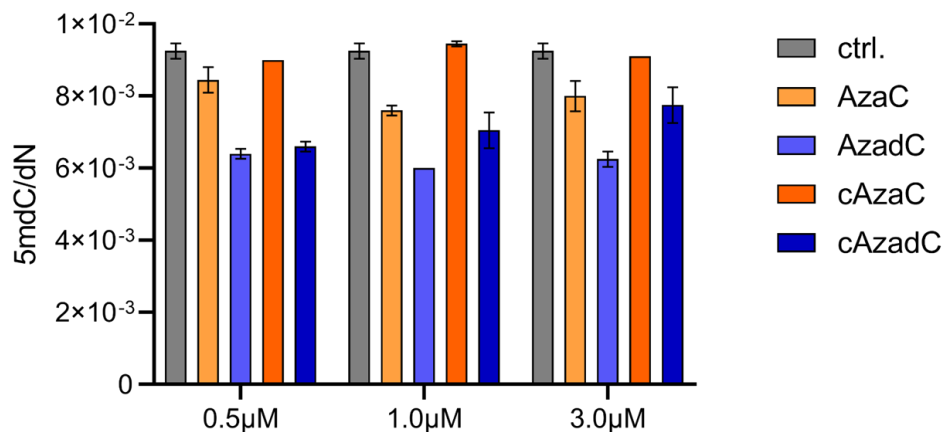


Abbildung 6-35: Ergebnisse der massenspektrometrischen Quantifizierung von mdC bei Zellverfütterung von AzaC, AzadC, cAzaC und cAzadC im Vergleich zur ungefütterten Kontrolle.

Zusammenfassend kann also festgestellt werden, dass das hier synthetisierte carbazyklische Derivat des 5-Azacytidins bei Zellverfütterungsexperimenten an die oben genannten Leukämie-Zelllinien keinen detektierbaren Effekt auf den Gehalt von mdC in der DNA der Zellen aufwies. Als eine mögliche Erklärung für die erhaltenen Ergebnisse könnte angeführt werden, dass die in der Zelle vorhandenen Kinasen das verfütterte cAzaC nicht als Substrat erkennen und zum Nukleotid phosphorylieren konnten, weswegen eine Umwandlung zum DNA-Nukleotid und ein Einbau in die DNA der Zellen nicht erfolgt sein konnte. Ein Scheitern der Zellverfütterungsexperimente basierend auf der Nichtaufnahme von cAzaC der Zellen in ihren löslichen Nukleosidpool schien unwahrscheinlich, da Zellverfütterungsexperimente mit dem analogen, Ribose-basierten 5-Aza-C zu einer signifikanten Reduktion des mdC- und hmdC-Gehalts in der DNA der untersuchten Zelllinien führten.

6.5.4. Synthese eines Nukleotid-*Prodrugs* von cAzaC und Untersuchung dessen in Zellverfütterungsexperimenten

Eine Möglichkeit den ersten enzymatischen Phosphorylierungsschritt im löslichen Nukleotidpool der gefütterten Zellen zu umgehen, stellt die Verfütterung eines Nukleotid-Monophosphat-*Prodrugs* dar.^[134,135] Dieses kann *in vivo* in das korrespondierende Monophosphat umgewandelt und im Anschluss durch die zellulären Kinasen weiter phosphoryliert, sowie möglicherweise zum DNA-Nukleotid prozessiert werden. Eine direkte Zellverfütterung des Monophosphat-Nukleotids kann somit umgangen werden, da die mehrfach negativ geladene Struktur des Monophosphats eine direkte Zellverfütterung auf Grund der äußerst geringen Permeabilität des Monophosphats durch die Zellmembran erschweren würde.

Nukleosid-Monophosphat-*Prodrugs* sind bereits gründlich erforscht und finden auch im klinischen Einsatz Anwendung. Ein weitbekanntes Nukleotid-*Prodrug* ist Sofosbuvir, dessen 2'-Fluoro-2'-desoxyuridin-Grundstruktur eine Phosphoroxymethylamidat-Einheit aufweist.^[136] Diese enthält einen L-Alanin-isopropylester, welcher *in vivo* in einem ersten Schritt durch zelluläre Esterasen gespalten werden kann. Anschließend erfolgt eine intramolekulare Zyklisierung unter Eliminierung von Phenol. Das zyklische Produkt kann schließlich hydrolysiert und abschließend durch spezifische Phosphoramidase-Enzyme gespalten werden, um das Monophosphat zu generieren.^[134]

Eine dem Sofosbuvir analoge Phosphoroxarylamidat-Einheit sollte im Rahmen dieses Projekts nun am cAzaC synthetisiert werden, um möglicherweise den ersten *in vivo* Phosphorylierungsschritt bei den Zellverfütterungsexperimenten dieses Monophosphat-Vorläufers zu umgehen. So könnte unter Umständen in der Konsequenz ein Effekt der Verbindung auf die DNA-Methylierung erzielt werden.

Die synthetische Einführung solcher Phosphoroxarylamidat-Einheiten an DNA- und RNA-Nukleosiden ist bereits umfänglich in der Literatur beschrieben,^[137] wobei die am weitesten verbreiteten Methoden eine Aktivierung der 5'-Hydroxygruppe ohne Einführung weiterer Schutzgruppen am Nukleosid vorsehen. Die unmittelbare Aktivierung wird dabei entweder durch die Zugabe von *N*-Methylimidazol (NMI) oder durch vorherige Deprotonierung der 5'-Hydroxygruppe durch *tert*-Butylmagnesiumchlorid erreicht.^[137] Im Rahmen dieses Projekts wurden diese beiden Methoden unter Verwendung verschiedener Reaktionsbedingungen für die Reaktion des Phosphoroxarylchloridats^[138,139] **81** mit cAzaC untersucht, wobei bei keinem Ansatz mehr als nur analytische Spuren des gewünschten Produkts nachweisbar waren.

Um die Reaktionsausbeute zu verbessern, wurde deswegen eine bereits in der Literatur für Ribonukleoside beschriebene Isopropylidenschutzgruppe an der 2'- und 3'-Position eingeführt.^[140] Durch Schützung dieser Hydroxygruppen sollten Nebenreaktionen bei der Einführung der Phosphoroxarylamidat-Einheit vermieden und somit die Bildung des gewünschten Produkts verstärkt werden. Verbindung **80** wurde durch Reaktion von cAzaC **79** mit 2,2-Dimethoxypropan unter Säurekatalyse und anschließender Neutralisation mit Triethylamin in annehmbarer Ausbeute erhalten. Anschließend konnte das geschützte Phosphoroxarylamidat **82** unter Aktivierung der 5'-Hydroxygruppe der Isopropylidinacetal-geschützten Verbindung **80** mit *tert*-Butylmagnesiumchlorid und anschließender Reaktion im Eintopfverfahren mit Phosphoroxarylchloridat **81** in einer Ausbeute von 67% nach säulenchromatographischer Aufreinigung synthetisiert werden (siehe Abbildung 6-36).

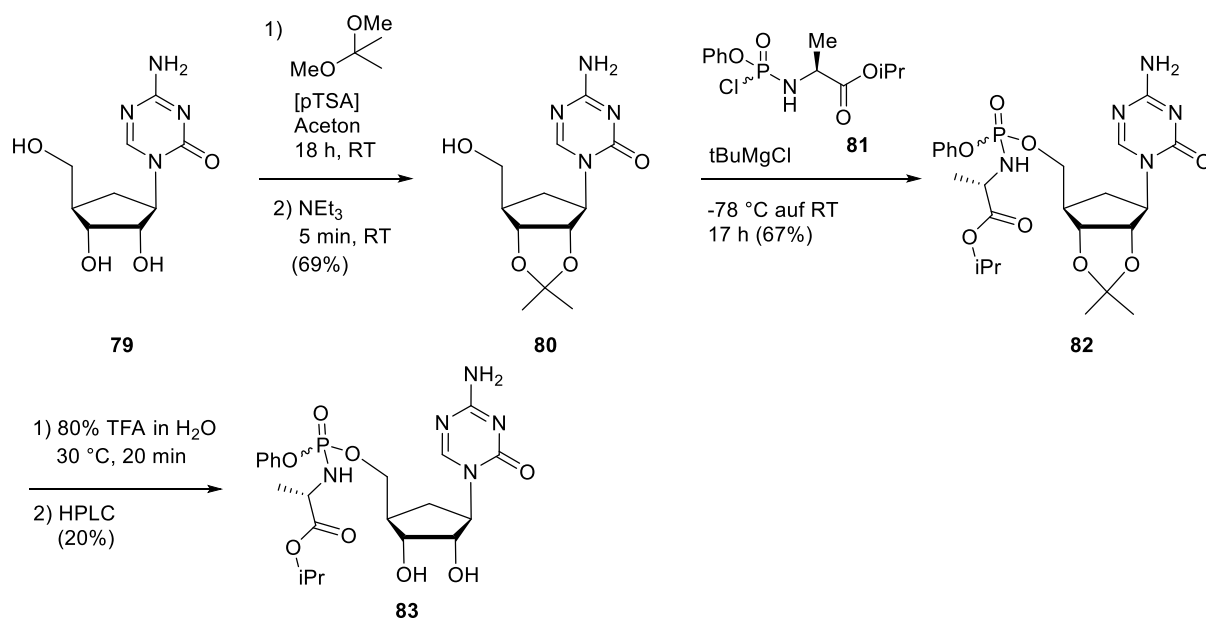


Abbildung 6-36: Schema der Synthese des Monophosphat-*Prodrugs* **83** von cAzaC.

Abschließend konnte Verbindung **82** unter stark sauren Bedingungen durch Verwendung von 80% wässriger Trifluoressigsäure entschützt und abschließend HPL-chromatographisch aufgereinigt werden. Abbildung 6-37 zeigt einen ausgewählten Bereich des erhaltenen HPL-Chromatogramms des Rohprodukts **83**, wobei zu erkennen ist, dass zwei Diastereomere am Phosphorzentrum erhalten wurden. Da es sich um diastereomere Verbindungen handelt, wurden diese selbstverständlich auch nicht im Verhältnis 1:1 erhalten.

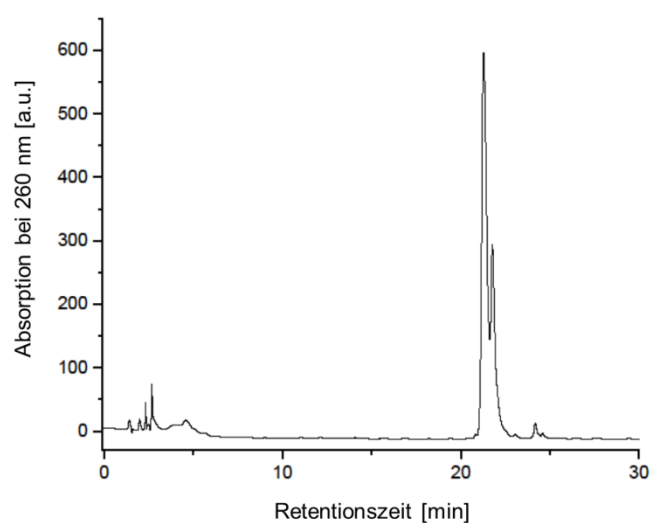


Abbildung 6-37: HPL-Chromatogramm des Rohprodukts **83** der Prodrugsynthese, welches die beiden diastereomeren Reaktionsprodukte zeigt.

Die Bildung von zwei Diastereomeren war zu erwarten, da die Reaktion von Verbindung **80** mit Phosphoroxarylchloridat **81** nicht stereospezifisch durchgeführt und somit bereits hier ein Gemisch aus zwei Diastereomeren am Phosphorzentrum erhalten wurde, welches säulenchromatographisch nicht aufgetrennt werden konnte. Nichtsdestotrotz war eine Auftrennung der beiden diastereomeren Verbindungen bei der HPL-chromatographischen Aufreinigung des Endprodukts weitestgehend möglich, weswegen in der Folge die biologische Wirksamkeit der beiden diastereomeren Verbindungen separat und im Gemisch untersucht werden konnte. Dies ist insofern relevant, da bereits in der Literatur beschrieben wird, dass die zellulären Esterasen, die die Isopropylesterereinheit des eingesetzten *Prodrugs* in einem ersten Schritt *in vivo* spalten sollten, unterschiedliche Affinitäten für die beiden Diastereomere am Phosphorzentrum aufweisen.^[140]

Im Anschluss wurden Zellverfütterungsexperimente der beiden diastereomeren *Prodrug*-Verbindungen des cAzaC separat sowie eines Gemisches der beiden Verbindungen an HEK293T-Zellen von Dr. Franziska Traube durchgeführt.

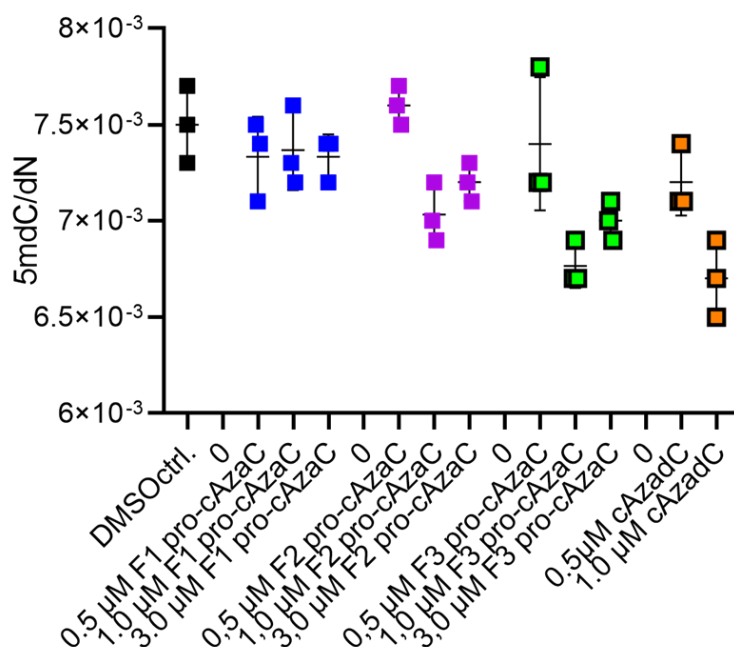


Abbildung 6-38: Massenspektrometrische Quantifizierung von mdC bei der Verfütterung der ersten eluierenden, diastereomeren Verbindung des *Prodrugs* (F1), einem Gemisch der beiden Diastereomere (F2) und der als zweites eluierenden, diastereomeren Verbindung des *Prodrugs* (F3) in jeweils verschiedenen Konzentrationen an HEK-Zellen im Vergleich zur ungefütterten Kontrolle; ein biologisches Replikate gemessen als technisches Triplikate.

Die HEK-Zellen wurden hier auf Grund ihrer einfacheren Kultivierbarkeit gewählt, im Falle erfolgreicher Experimente könnten diese dann an den oben beschriebenen Leukämie-Zelllinien durchgeführt werden. Abbildung 6-38 zeigt die Ergebnisse der massenspektrometrischen Analyse der aus den Zellverfütterungsexperimenten erhaltenen DNA, wobei zur Kontrolle ungefütterte Proben sowie mit cAza-dC behandelte Proben ebenfalls untersucht wurden.

Abbildung 6-38 zeigt deutlich, dass insbesondere die als zweites eluierende diastereomere Verbindung einen deutlich reduzierenden Effekt auf den mdC Gehalt der DNA bei Zellverfütterung an HEK-Zellen im Vergleich zur ungefütterten Kontrollprobe aufweist. Dieser ist bei Verfütterung der Verbindung in Konzentrationen von 1.0 μM und 3.0 μM vergleichbar mit dem hypomethylierenden Effekt, der bei Verfütterung von 1.0 μM cAza-dC erreicht wird. Auch das Gemisch der beiden diastereomeren Verbindungen weist einen leicht hypomethylierenden Effekt auf, was auch damit begründet werden kann, dass die zuerst eluierende diastereomere Verbindung keinen nachweisbaren Effekt auf den mdC-Gehalt der DNA bei den durchgeführten Experimenten erzielte, weswegen ein Gemisch aus einer hypomethylierenden Verbindung und einer inaktiven Verbindung zu einem abgeschwächten Ergebnis im Vergleich zur Verfütterung der aktiven Reinsubstanz führt.

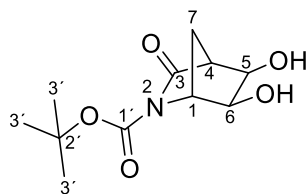
Zusammenfassend kann also festgestellt werden, dass durch die Synthese eines Monophosphat-*Prodrugs* ein carbazyklisches Derivat des 5-Aza-C hergestellt werden konnte, das bei Zellverfütterungsexperimenten an HEK-Zellen den gewünschten hypomethylierenden Effekt unter gleichzeitiger Hydrolysebeständigkeit zeigte.

6.5.5. Projektbeiträge

Im Rahmen dieses Projekts wurden von mir die Synthese des carbazyklischen 5-Azacytidins, die Synthese des Nukleotid-*Prodrugs* dieser Verbindung sowie die Stabilitätsuntersuchungen in wässriger Umgebung entwickelt und durchgeführt. Die Zellverfütterungsexperimente wurden von *Dr. Franziska Traube*, die massenspektrometrischen Messungen sowie die anschließende Datenauswertung wurden von *Florian Schelter* in Zusammenarbeit mit *Dr. Franziska Traube* durchgeführt.

6.5.6. Experimentalteil

***tert*-Butyl-(1R,4S,5R,6S)-5,6-dihydroxy-3-oxo-2-azabicyclo[2.2.1]heptan-2-carboxylat (71)**



Das Boc-geschützte *Vince*-Lactam^[32] **70** (11.66 g, 55.7 mmol, 1.0 eq.) wurde in 214 mL THF, 92 mL *tert*-Butanol und 31 mL H₂O gelöst, mit 22.59 g *N*-Methylmorpholino-*N*-oxid (167.2 mmol, 3.0 eq.) and 1.03 g K₂OsO₄-Dihydrat (2.79 mmol, 5 mol%) versetzt und das Reaktionsgemisch wurde 16 h bei Raumtemperatur verrührt. Anschließend wurden 200 mL Wasser und 35.1 g Na₂SO₃ (278.5 mmol, 5.0 eq.) zugegeben und das Gemisch wurde weitere 1 h bei Raumtemperatur verrührt. Daraufhin wurde das Gemisch mit EtOAc (3 x 300 mL) extrahiert, die vereinigten organischen Phasen wurden mit gesättigter wässriger Natriumchloridlösung (300 mL) gewaschen und über Na₂SO₄ getrocknet. Nach Entfernen der Lösungsmittel *in vacuo* wurde der Rückstand mittels Säulenchromatographie an Silicagel aufgereinigt (CH₂Cl₂:MeOH 40:1 → CH₂Cl₂:MeOH 30:1), um 2.91 g des *exo*-bishydroxylierten Produktes **71** (12.0 mmol, 22 %) als gelben Schaum zu liefern.

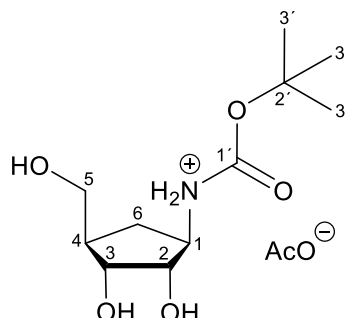
¹H-NMR (400 MHz, CDCl₃): δ/ppm = 4.35 – 4.32 (m, 1H, 1-H), 4.23 (d, *J* = 6.0 Hz, 1H, 5-H), 4.08 (d, *J* = 6.1 Hz, 1H, 6-H), 3.89 – 3.80 (br s, -OH), 3.78 – 3.68 (br s, -OH), 2.82 – 2.76 (m, 1H, 4-H), 2.09 (dt, *J* = 10.9, 1.5 Hz, 1H, 7-H), 1.97 (m, 1H, 7-H), 1.51 (s, 9H, 3'-H).

¹³C-NMR (101 MHz, CDCl₃): δ/ppm = 173.30 (3-C), 149.20 (1'-C), 83.82 (2'-C), 70.57 (6-C), 68.14 (5-C), 62.29 (1-C), 53.74 (4-C), 31.99 (7-C), 28.16 (3'-C).

IR (ATR): ν (cm⁻¹) = 3415 (w), 2979 (w), 1777 (m), 1714 (m), 1457 (w), 1394 (w), 1368 (m), 1350 (m), 1326 (m), 1297 (m), 1256 (m), 1144 (s), 1081 (s), 1043 (w), 1031 (w), 1016 (m), 947 (s), 902 (w), 838 (w), 776 (m), 733 (s), 702 (m).

R_f (CH₂Cl₂:MeOH 20:1): 0.24.

**tert-Butyl-((1R,2S,3R,4R)-2,3-dihydroxy-4-(hydroxymethyl)-cyclopentyl)-
carbamate-Essigsäurekomplex (72)**



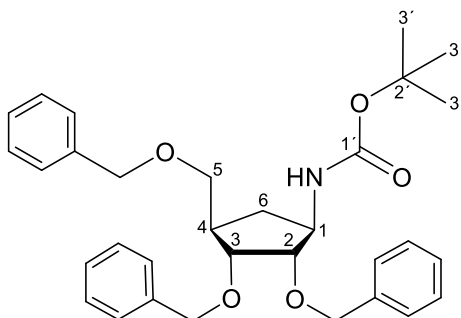
Eine Lösung des bishydroxylierten Lactams **71** (2.90 g, 11.9 mmol, 1.0 eq.) in 40 mL trockenem MeOH wurde auf 0 °C gekühlt und mit NaBH₄ (2.25 g, 59.61 mmol, 5.0 eq.) in 5 Portionen im Verlauf von 15 min versetzt. Anschließend wurde das Reaktionsgemisch 45 min bei 0 °C verrührt, auf Raumtemperatur erwärmen gelassen und weitere 1 h bei Raumtemperatur verrührt. Daraufhin wurden 100 mL verdünnte Essigsäure (10 % v/v in MeOH) und Silicagel zugegeben und das Gemisch wurde *in vacuo* zur Trockene eingeeengt. Aufreinigung des Rohprodukts mittels Säulenchromatographie an Silicagel (*dry load*, CH₂Cl₂:MeOH 5:1 → CH₂Cl₂:MeOH 4:1) lieferte nach Coevaporation der produkthaltigen Fraktionen mit trockenem Toluol (2 x 150 mL) das Reduktionsprodukt **72** in Form des Acetatsalzes (2.96 g, 9.6 mmol, 81 %) als bräunliches Öl.

¹H-NMR (400 MHz, DMSO-*d*₆): δ/ppm = 6.70 (d, *J* = 7.9 Hz, 1H, NH), 3.65 – 3.49 (m, 2H, 1-H, 3-H), 3.52 (dd, *J* = 6.3, 5.2 Hz, 1H, 2-H), 3.37 (dd, *J* = 10.5, 5.8 Hz, 1H, 5-H), 3.30 (dd, *J* = 10.5, 6.2 Hz, 1H, 5-H), 1.98 (dt, *J* = 13.1, 8.4 Hz, 1H, 6-H), 1.86 – 1.77 (m, 4H, 4-H, OAc), 1.37 (s, 9H, 3'-H), 1.04 – 0.90 (m, 1H, 6-H).

¹³C-NMR (101 MHz, DMSO-*d*₆): δ/ppm = 173.28 (OAc-C1), 155.23 (C-1'), 77.37 (2'-C), 75.93 (2-C), 72.07 (1-C), 63.02 (5-C), 55.07 (3-C), 44.94 (4-C), 30.61 (6-C), 28.30 (3'-C), 22.49 (OAc-C2).

R_f (CH₂Cl₂:MeOH 8:1): 0.43.

tert-Butyl-((1R,2S,3R,4R)-2,3-bis(benzyloxy)-4-((benzyloxy)methyl)-cyclopentyl)-carbamat (73)



Eine Lösung des Cyclopentans **72** (2.93 g, 9.5 mmol, 1.0 eq.) in 67 mL trockenem DMF wurde auf 0 °C gekühlt und mit NaH (1.57 g, 39.3 mmol, 4.1 eq., 60 % in Mineralöl) in drei Portionen versetzt. Das Gemisch wurde anschließend 15 min bei 0 °C verrührt, daraufhin tropfenweise mit 5.31 mL BnBr (44.6 mmol, 4.7 eq.) versetzt und 2.5 h bei 0 °C und weiter 3 h bei Raumtemperatur verrührt. Im Anschluss wurden 220 mg TBAI (0.6 mmol, 6 mol%) und weitere 3.18 mL BnBr (26.8 mmol, 2.8 eq.) zugegeben und das Gemisch wurde auf 80 °C erwärmt und 3.5 h bei dieser Temperatur verrührt. Schließlich wurde das Gemisch auf Raumtemperatur abkühlen gelassen, in 300 mL gesättigte wässrige NaHCO₃-Lösung gegossen und mit CH₂Cl₂ extrahiert (3 x 300 mL). Die vereinigten organischen Phasen wurden mit gesättigter wässriger NaHCO₃-Lösung (300 mL) und gesättigter wässriger Natriumchloridlösung (300 mL) gewaschen, über wasserfreiem Na₂SO₄ getrocknet und *in vacuo* bei 50 °C eingeeengt. Der Rückstand wurde mittels Säulenchromatographie an Silicagel aufgereinigt (iHex:EtOAc 10:1 → iHex:EtOAc 8:1 → iHex:EtOAc 7:1 → iHex:EtOAc 6:1 → iHex:EtOAc 4:1), um das dreifach Benzyl-geschützte Produkt **73** (2.47 g, 4.8 mmol, 51 %) als farbloses, festes Wachs zu liefern.

¹H-NMR (400 MHz, CDCl₃): δ/ppm = 7.45 – 7.09 (m, 15H, 3 x OCH₂Ph), 4.95 (s, 1H, NH), 4.76 – 4.25 (m, 6H, 3 x OCH₂Ph), 4.15 – 3.90 (m, 1H, 1-H), 3.82 (dd, *J* = 6.7, 4.4 Hz, 1H, 3-H), 3.79 – 3.63 (m, 1H, 2-H), 3.57 – 3.34 (m, 2H, 5-H), 2.51 – 2.24 (m, 2H, 4-H, 6-H), 1.50 – 1.34 (m, 9H, 3'-H), 1.31 – 1.19 (m, 1H, 6-H).

¹³C-NMR (101 MHz, CDCl₃): δ/ppm = 155.09 (1'-C), 138.48 (Ar-C), 138.43 (Ar-C), 138.13 (Ar-C), 128.44 (Ar-C), 128.30 (Ar-C), 128.24 (Ar-C), 128.03 (Ar-C), 127.88 (Ar-C), 127.82 (Ar-C), 127.70 (Ar-C), 127.54 (Ar-C), 127.51 (Ar-C), 81.27 (2-C),

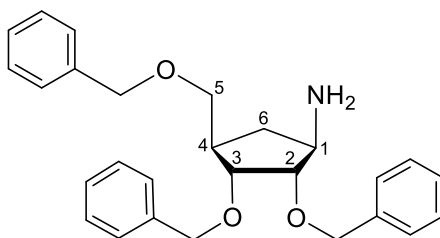
79.29 (3-C), 79.13 (2'-C), 73.30 (OCH₂-Ph), 71.52 (OCH₂-Ph), 71.02 (OCH₂-Ph), 70.41 (5-C), 52.85 (1-C), 41.49 (1-C), 30.34 (6-C), 28.46 (3'-C).

HRMS (ESI⁺): m/z berechnet für C₃₂H₄₀NO₅⁺ [M+H]⁺ 518.2901; gefunden: 518.2911.

IR (ATR): ν (cm⁻¹) = 3331 (br w), 3029 (w), 2974 (w), 2867 (w), 1705 (s), 1495 (m), 1453 (m), 1390 (w), 1364 (s), 1245 (m), 1165 (s), 1092 (s), 1071 (s), 1027 (m), 732 (s), 695 (s).

R_f (iHex:EtOAc 4:1): 0.69.

**(1R,2S,3R,4R)-2,3-Bis(benzyloxy)-4-((benzyloxy)methyl)-cyclopentan-1-amin
(74)**



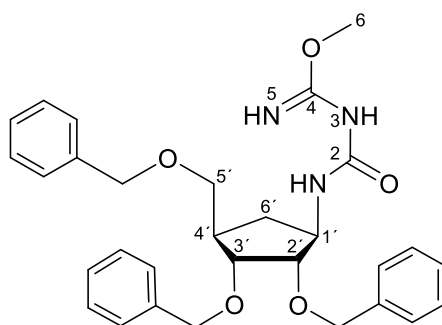
Unter Ar-Atmosphäre wurden 1.02 g des Boc-geschützten Cyclopentans **73** (1.97 mmol, 1.0 eq.) in 7.0 mL trockenem CH₂Cl₂ gelöst, daraufhin wurden 3.0 mL TFA zugetropft und das Reaktionsgemisch 1 h bei Raumtemperatur verrührt. Die erhaltene rote Lösung wurde zur Trockene eingeeengt, mit 20.0 mL gesättigter wässriger NaHCO₃-Lösung versetzt und das Gemisch wurde 10 min bei Raumtemperatur verrührt. Anschließend wurden 60 mL EtOAc zugegeben, die Phasen wurden getrennt und die wässrige Phase wurde weiter mit EtOAc (2 x 60 mL) extrahiert. Die vereinigten organischen Phasen wurden über Na₂SO₄ getrocknet und *in vacuo* zur Trockene eingeeengt, um die an der Aminfunktion entschützte Verbindung **74** in Form eines gelben Öls zu liefern (882 mg, 1.97 mmol, quant.), welche im nächsten Schritt ohne weitere Aufreinigung auf Grund ihrer Oxidationsanfälligkeit weiterverwendet wurde.

¹H-NMR (400 MHz, CDCl₃): δ /ppm = 7.39 – 7.26 (m, 15H, 3 x OCH₂Ph), 4.63 – 4.33 (m, 6H, 3 x OCH₂Ph), 3.81 (dd, *J* = 5.0, 3.1 Hz, 1H, 3-H), 3.49 (dt, *J* = 9.1, 7.8 Hz, 1H, 1-H), 3.41 (dd, *J* = 9.2, 5.1 Hz, 1H, 5-H), 3.34 – 3.26 (m, 2H, 2-H, 5-H), 2.47 – 2.34 (m, 1H, 4-H), 2.24 – 2.13 (m, 1H, 6-H), 1.06 – 0.94 (m, 1H, 6-H).

¹³C-NMR (101 MHz, CDCl₃): δ/ppm = 138.66 (Ar-C), 138.52 (Ar-C), 138.48 (Ar-C), 128.72 (Ar-C), 128.52 (Ar-C), 128.47 (Ar-C), 128.38 (Ar-C), 128.31 (Ar-C), 128.00 (Ar-C), 127.74 (Ar-C), 127.67 (Ar-C), 127.13 (Ar-C), 86.10 (2-C), 78.11 (3-C), 73.21 (OCH₂-Ph), 72.27 (5-C), 71.86 (OCH₂-Ph), 70.89 (OCH₂-Ph), 54.32 (1-C), 41.49 (4-C), 32.19 (6-C).

HRMS (ESI⁺): m/z berechnet für C₂₇H₃₂NO₃⁺ [M+H]⁺ 418.2377; gefunden: 418.2378.

1-[(1'R,2'S,3'R,4'R)-2',3'-Bis(benzyloxy)-4'-((benzyloxy)methyl)-cyclopentyl]-methylisobiuret (76)



Unter Argonatmosphäre wurden 882 mg des Cyclopentylamins **74** (1.87 mmol, 1.97 mmol, 1.0 eq.) in 18.20 mL trockenem MeCN gelöst und mit 431 mg 2-Methyl-1-(1-imidazolylcarbonyl)-isoharnstoff^[32] **75** (2.56 mmol, 1.3 eq.) versetzt. Das Gemisch wurde anschließend 2 h bei 92 °C refluxiert, auf Raumtemperatur abkühlen gelassen und daraufhin zur Trockene eingeeengt. Der Rückstand wurde mittels Säulenchromatographie an Silicagel aufgereinigt (iHex:EtOAc 3:1 → iHex:EtOAc 5:2 → iHex:EtOAc 2:1 → iHex:EtOAc 1:1), um Methylisobiuret **76** als farbloses Öl zu erhalten (886 mg, 1.17 mmol, 87 %).

¹H-NMR (400 MHz, CDCl₃): δ/ppm = 7.43 – 7.22 (m, 15H, 3 x OCH₂Ph), 5.60 (s, 1H, 1-H), 4.80 – 4.61 (m, 2H, OCH₂Ph), 4.55 – 4.41 (m, 3H, OCH₂Ph), 4.32 (d, *J* = 11.8 Hz, 1H, OCH₂Ph), 4.24 (td, *J* = 7.6, 3.4 Hz, 1H, 1'-H), 3.87 (dd, *J* = 6.9, 4.5 Hz, 1H, 3'-H), 3.82 (dd, *J* = 4.5, 3.1 Hz, 1H, 2'-H), 3.61 (s, 2H, 3-H, 5-H), 3.57 – 3.45 (m, 2H, 5'-H), 2.54 – 2.39 (m, 2H, 4'-H, 6'-H), 2.17 (s, 3H, 6-H), 1.39 – 1.31 (m, 1H, 6'-H).

¹³C-NMR (101 MHz, CDCl₃): δ/ppm = 207.11 (4-C), 162.27 (2-C), 138.74 (Ar-C), 138.71 (Ar-C), 138.45 (Ar-C), 128.51 (Ar-C), 128.34 (Ar-C), 128.28 (Ar-C), 127.86 (Ar-C), 127.70 (Ar-C), 127.59 (Ar-C), 127.56 (Ar-C), 81.65 (2'-C), 79.79 (3'-C), 73.35

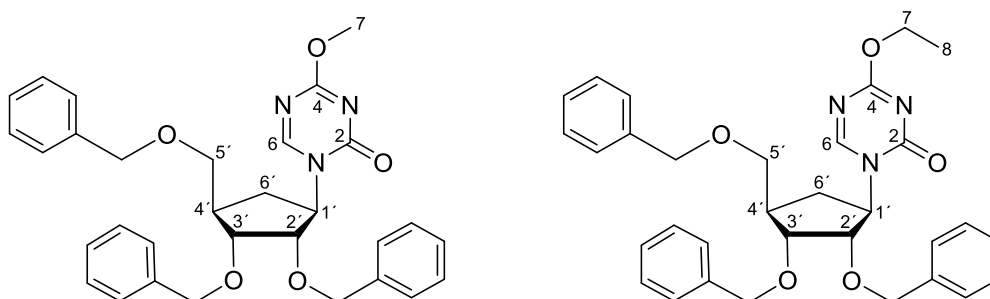
(OCH₂-Ph), 71.64 (OCH₂-Ph), 71.20 (OCH₂-Ph), 70.60 (5'-C), 52.56 (1'-C), 41.68 (4'-C), 31.07 (6-C), 30.57 (6'-C).

HRMS (ESI⁺): berechnet für C₃₀H₃₆N₃O₅⁺ [M+H]⁺ 518.2650; gefunden: 518.2652.

IR (ATR): ν (cm⁻¹) = 3384 (br w), 3028 (w), 2857 (w), 1694 (w), 1635 (s), 1495 (s), 1453 (m), 1364 (m), 1300 (w), 1197 (w), 1096 (s), 1027 (m) 799 (w), 736 (m), 697 (s).

R_f (iHex:EtOAc 2:1): 0.37.

4-Methoxy-1-[(1'*R*,2'*S*,3'*R*,4'*R*)-2',3'-bis(benzyloxy)-4'-((benzyloxy)methyl)-cyclopentyl]-1H-[1,3,5]-triazin-2-on (77a) und 4-Ethoxy-1-[(1'*R*,2'*S*,3'*R*,4'*R*)-2',3'-bis(benzyloxy)-4'-((benzyloxy)methyl)-cyclopentyl]-1H-[1,3,5]-triazin-2-on (77b)



Methylisobiuret **76** (837 mg, 1.62 mmol, 1.0 eq.) wurde in 15.30 mL Triethylorthoformiat gelöst und mit 20 μ L TFA versetzt. Das Gemisch wurde 3 h auf 90 °C, anschließend auf Raumtemperatur abkühlen gelassen und *in vacuo* bei 50 °C zur Trockene eingeeengt. Der Rückstand wurde mit trockenem MeOH (30 mL) coevaporiert und mittels Säulenchromatographie an Silicagel aufgereinigt (iHex:EtOAc 3:1 \rightarrow iHex:EtOAc 5:2 \rightarrow iHex:EtOAc 2:1 \rightarrow iHex:EtOAc 1:1 \rightarrow iHex:EtOAc 1:2), um 373 mg des Ethoxytriazins **77b** (0.69 mmol, 43 %) in Form eines farblosen Wachses und 389 mg des Methoxytriazins **77a** (0.73 mmol, 46 %) in Form eines farblosen Feststoffs zu erhalten.

Methoxytriazin **77a**:

¹H-NMR (400 MHz, CDCl₃): δ /ppm = 7.96 (s, 1H, 6-H), 7.38 – 7.17 (m, 15H, 3 x OCH₂Ph), 4.70 – 4.40 (m, 6H, 1'-H, 2.5 x OCH₂Ph), 4.36 – 4.24 (m, 2H, 2'-H, 0.5 x OCH₂Ph), 3.98 (s, 3H, 7-H), 3.93 (dd, *J* = 4.9, 3.1 Hz, 1H, 3'-H), 3.54 – 3.40 (m, 2H, 5'-H), 2.51 (ddt, *J* = 9.8, 4.4, 2.3 Hz, 1H, 4'-H), 2.35 (dt, *J* = 13.3, 9.4 Hz, 1H, 6'-H), 1.82 (ddd, *J* = 13.4, 9.0, 6.7 Hz, 1H, 6'-H).

^{13}C -NMR (101 MHz, CDCl_3): δ/ppm = 169.87 (4-C), 159.59 (6-C), 154.67 (2-C), 138.16 (Ar-C), 138.14 (Ar-C), 137.63 (Ar-C), 128.63 (Ar-C), 128.60 (Ar-C), 128.51 (Ar-C), 128.24 (Ar-C), 128.20 (Ar-C), 128.17 (Ar-C), 127.90 (Ar-C), 127.86 (Ar-C), 78.78 (2'-C), 77.27 (3'-C), 73.40 ($\text{OCH}_2\text{-Ph}$), 72.17 ($\text{OCH}_2\text{-Ph}$), 71.41 ($\text{OCH}_2\text{-Ph}$), 71.20 (5'-C), 63.34 (1'-C), 55.95 (7-C), 41.14 (4'-C), 26.75 (6'-C).

HRMS (ESI⁺): m/z berechnet für $\text{C}_{31}\text{H}_{34}\text{N}_3\text{O}_5^+$ $[\text{M}+\text{H}]^+$ 528.2493; gefunden: 528.2498.

IR (ATR): ν (cm^{-1}) = 3029 (w), 2867 (w), 1698 (s), 1613 (s), 1515 (m), 1464 (m), 1453 (m), 1398 (m), 1337 (m), 1265 (w), 1211 (w), 1096 (m), 1070 (m), 1027 (m), 913 (w), 802 (m), 733 (s), 696 (s).

R_f (iHex:EtOAc 2:1): 0.27.

Ethoxytriazin 77b:

^1H -NMR (400 MHz, CDCl_3): δ/ppm = 7.97 (s, 1H, 6-H), 7.39 – 7.15 (m, 15H, 3 x OCH_2Ph), 4.66 – 4.21 (m, 10H, 1'-H, 2'-H, 7-H, 3 x OCH_2Ph), 3.96 – 3.87 (m, 1H, 3'-H), , 3.54 – 3.36 (m, 2H, 5'-H), 2.56 – 2.41 (m, 1H, 4'-H), 2.33 (ddd, J = 14.5, 9.5, 4.7 Hz, 1H, 6'-H), 1.88 – 1.74 (m, 1H, 6'-H), 1.25 (q, J = 6.9 Hz, 3H, 8-H).

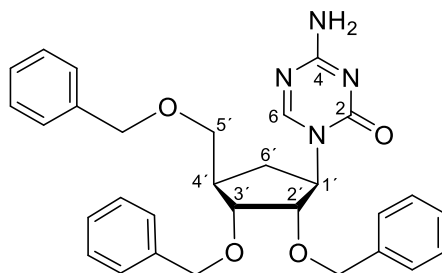
^{13}C -NMR (101 MHz, CDCl_3): δ/ppm = 169.31 (4-C), 159.53 (6-C), 154.74 (2-C), 138.18 (Ar-C), 138.16 (Ar-C), 137.64 (Ar-C), 128.79 (Ar-C), 128.64 (Ar-C), 128.61 (Ar-C), 128.51 (Ar-C), 128.25 (Ar-C), 128.20 (Ar-C), 128.19 (Ar-C), 127.90 (Ar-C), 127.87 (Ar-C), 79.11 (2'-C), 77.36 (3'-C), 73.46 ($\text{OCH}_2\text{-Ph}$), 72.28 ($\text{OCH}_2\text{-Ph}$), 71.55 ($\text{OCH}_2\text{-Ph}$), 71.20 (5'-C), 65.06 (7-C), 63.33 (1'-C), 41.15 (4'-C), 26.81 (6'-C), 14.24 (8-C).

HRMS (ESI⁺): m/z berechnet für $\text{C}_{32}\text{H}_{36}\text{N}_3\text{O}_5^+$ $[\text{M}+\text{H}]^+$ 542.2650; gefunden: 542.2656.

IR (ATR): ν (cm^{-1}) = 3064 (w), 3030 (w), 2924 (w), 2962 (w), 1694 (s), 1615 (s), 1514 (m), 1495 (m), 1452 (s), 1361 (m), 1326 (m), 1270 (m), 1206 (w), 1096 (s), 1070 (s), 1027 (m), 803 (w), 734 (8s), 697 (s).

R_f (iHex:EtOAc 2:1): 0.30.

4-Amino-1-[(1'R,2'S,3'R,4'R)-2',3'-bis(benzyloxy)-4'-((benzyloxy)methyl)-cyclopentyl]-1H-[1,3,5]-triazin-2-on (78)



Ein Gemisch aus Ethoxytriazin **77b** (358 mg, 0.66 mmol) und Methoxytriazin **77a** (363 mg, 0.69 mmol) wurde in 26.8 mL NH_3 (7 N in MeOH) gelöst und 3 h bei Raumtemperatur verrührt. Daraufhin wurden dem Reaktionsgemisch 125 mL H_2O zugesetzt und schließlich mit EtOAc (3 x 200 mL) extrahiert. Die vereinigten organischen Phasen wurden über Na_2SO_4 getrocknet und *in vacuo* zur Trockene eingengt, um 593 mg des Aminierungsproduktes **78** (1.16 mmol, 86 %) in Form eines farblosen Schaums zu erhalten, der ohne weitere Aufreinigung für die nächste Stufe eingesetzt werden konnte.

$^1\text{H-NMR}$ (400 MHz, CDCl_3): δ/ppm = 7.80 (s, 1H, 6-H), 7.26 (m, 15H, 3 x OCH_2Ph), 5.70 (s, 1H, NH), 4.63 – 4.42 (m, 6H, 1'-H, 2.5 x OCH_2Ph), 4.39 – 4.28 (m, 2H, 2'-H, 0.5 x OCH_2Ph), 3.91 (dd, J = 5.0, 3.5 Hz, 1H, 3'-H), 3.53 – 3.42 (m, 2H, 5'-H), 2.59 – 2.35 (m, 1H, 4'-H), 2.26 (dt, J = 13.3, 9.3 Hz, 1H, 6'-H), 1.86 (ddd, J = 13.3, 9.2, 7.1 Hz, 1H, 6'-H)

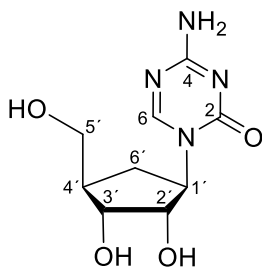
$^{13}\text{C-NMR}$ (101 MHz, CDCl_3): δ/ppm = 165.84 (4-C), 158.30 (6-C), 154.11 (2-C), 138.27 (Ar-C), 138.24 (Ar-C), 137.83 (Ar-C), 128.57 (Ar-C), 128.56 (Ar-C), 128.47 (Ar-C), 128.21 (Ar-C), 128.12 (Ar-C), 128.05 (Ar-C), 127.84 (Ar-C), 127.82 (Ar-C), 127.78 (Ar-C), 78.99 (2'-C), 77.39 (3'-C), 73.30 ($\text{OCH}_2\text{-Ph}$), 72.18 ($\text{OCH}_2\text{-Ph}$), 71.39 ($\text{OCH}_2\text{-Ph}$), 71.3 (5'-C), 63.53 (1'-C), 41.21 (4'-C), 26.79 (6'-C).

HRMS (ESI⁺): m/z berechnet für $\text{C}_{30}\text{H}_{33}\text{N}_4\text{O}_4^+$ $[\text{M}+\text{H}]^+$ 513.2496; gefunden: 513.2496.

IR (ATR): ν (cm^{-1}) = 3331 (br w), 3182 (br w), 3060 (w), 3028 (w), 2925 (w), 2857 (w), 1633 (s), 1506 (m), 1495 (8m), 1470 (m), 1452 (8s), 1361 (w), 1262 (w), 1092 (s), 1070 (s), 1027 (m), 912 (w), 799 (m), 734 (s), 696 (s).

R_f ($\text{CH}_2\text{Cl}_2\text{:MeOH}$ 15:1): 0.42.

4-Amino-1-[(1'R,2'S,3'R,4'R)-2',3'-bishydroxy-4'-(hydroxymethyl)-cyclopentyl]-1H-[1,3,5]-triazin-2-on (79)



Eine Lösung des Benzyl-geschützten Derivats von 5-Aza-carbaC **78** (576 mg, 1.12 mmol, 1.0 eq.) wurde in 30.8 mL trockenem CH₂Cl₂ unter Argonatmosphäre gelöst und auf -78 °C gekühlt. Daraufhin wurden 5.88 mL BCl₃ (1 M in CH₂Cl₂, 5.88 mmol, 5.25 eq.) zugetropft und das Reaktionsgemisch wurde 1 h bei -78 °C und 2 h bei Raumtemperatur verrührt. Daraufhin wurden 40 mL trockenes MeOH zugesetzt, das Reaktionsgemisch weitere 20 min bei Raumtemperatur verrührt und schließlich *in vacuo* zur Trockene eingengt. Der Rückstand wurde mittels Säulenchromatographie an Silicagel aufgereinigt (*dry load*, CH₂Cl₂:MeOH 8:1 → CH₂Cl₂:MeOH 5:1 → CH₂Cl₂:MeOH 4:1 → CH₂Cl₂:MeOH 7:3), um 228 mg der vollständig entschützten Verbindung **79** (0.94 mmol, 84 %) in Form eines farblosen Feststoffs zu erhalten.

Für die Anwendung in der Zellkultur wurden 34.5 mg der Verbindung weiter mittels *reversed phase* HPLC (0% auf 3% MeCN in H₂O in 45 min, Fluss 5 mL/min) aufgereinigt, um 25.3 mg der HPLC-aufgereinigten Verbindung **79** (73 %) zu erhalten.

¹H-NMR (400 MHz, D₂O): δ/ppm = 8.32 (s, 1H, 6-H), 4.57 – 4.33 (m, 2H, 1'-H, 2'-H), 4.04 (dd, *J* = 5.3, 4.4 Hz, 1H, 3'-H), 3.76 – 3.57 (m, 2H, 5'-H), 2.29 (dt, *J* = 13.0, 8.2 Hz, 1H, 6'-H), 2.23 -2.15 (m, 1H, 4'-H), 1.76 – 1.59 (m, 1H, 6'-H).

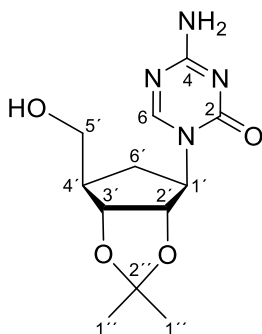
¹³C-NMR (101 MHz, D₂O): δ/ppm = 165.45 (4-C), 159.03 (6-C), 156.15 (2-C), 73.12 (2'-C), 71.77 (3'-C), 63.55 (1'-C), 62.90 (5'-C), 44.42 (4'-C), 27.17 (6'-C).

HRMS (ESI⁺): *m/z* berechnet für C₉H₁₅N₄O₄⁺ [M+H]⁺ 243.1088; gefunden: 243.1086.

HRMS (ESI⁻): *m/z* berechnet für C₉H₁₃N₄O₄⁻ [M-H]⁻ 241.0942; gefunden: 241.0941.

R_f (CH₂Cl₂:MeOH 7:3): 0.29.

4-Amino-1-((3a*S*,4*R*,6*R*,6a*R*)-6-(hydroxymethyl)-2,2-dimethyltetrahydro-4H-cyclopenta-[d][1,3]dioxol-4-yl)-1,3,5-triazin-2(1H)-on (80**)**



In einem ausgeheizten Kolben wurde 5-Aza-carbaC **79** (64 mg, 264 μ mol, 1.0 eq.) in 6.60 mL trockenem Aceton suspendiert. Anschließend wurde *p*-Toluolsulfonsäure (5 mg, katalytische Menge) und 2,2-Dimethoxypropan (49 μ L, 396 μ mol, 1.5 eq.) zugegeben und das Gemisch wurde 18 h bei Raumtemperatur verrührt. Zu der erhaltenen Lösung wurden 150 μ L NEt₃ (1.08 mmol) zugetropft und das Gemisch wurde weitere 10 min bei Raumtemperatur verrührt. Daraufhin wurde das Gemisch zur Trockene eingeeengt und der Rückstand mittels Säulenchromatographie an Silicagel aufgereinigt (CH₂Cl₂:MeOH 10:1, mit 1% NEt₃), um 81 mg der Isopropylidenacetal-geschützten Verbindung **80** (182 μ mol, 69%, Addukt mit 1.6 eq. NEt₃) in Form eines farblosen Feststoffs zu erhalten.

¹H-NMR (400 MHz, D₃COD): δ /ppm = 8.29 (s, 1H, 6-H), 4.94 (dd, *J* = 7.0, 4.9 Hz, 1H, 2'-H), 4.62 – 4.39 (m, 2H, 1'-H, 3'-H), 3.73 – 3.51 (m, 2H, 5'-H), 3.21 (q, *J* = 7.3 Hz, 10H, N(CH₂CH₃)₃), 2.32 – 2.07 (m, 3H, 4'-H, 6'-H), 1.50 (s, 3H, 1''-H), 1.32 (t, *J* = 7.3 Hz, 15H, N(CH₂CH₃)₃), 1.29 (s, 3H, 1''-H).

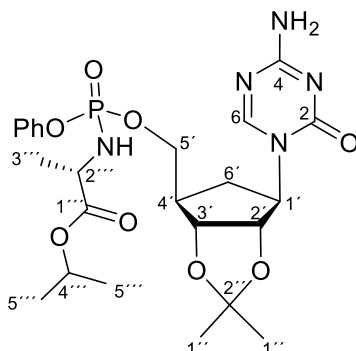
¹³C-NMR (101 MHz, D₃COD): δ /ppm = 167.87 (4-C), 159.88 (6-C), 156.66 (2-C), 114.31 (2''-C), 83.95 (2'-C), 82.98 (3'-C), 66.26 (1'-C), 64.17 (5'-C), 47.88 (N(CH₂CH₃)₃), 47.81 (4'-C), 33.72 (6'-C), 27.97 (1''-C), 25.53 (1''-C), 9.23 (N(CH₂CH₃)₃).

HRMS (ESI⁺): *m/z* berechnet für C₁₂H₁₉N₄O₄⁺ 283.1401 [M+H]⁺; gefunden: 283.1397.

IR (ATR): ν (cm⁻¹) = 3346 (br w), 2983 (w), 2658 (w), 1634 (s), 1472 (s), 1210 (m), 1160 (m), 1065 (m), 868 (w), 800 (m), 682 (w).

R_f (CH₂Cl₂:MeOH 6:1): 0.36.

Isopropyl((((3aR,4R,6R,6aS)-6-(4-amino-2-oxo-1,3,5-triazin-1(2H)-yl)-2,2-dimethyltetrahydro-4H-cyclopenta[d][1,3]dioxol-4-yl)methoxy)(phenoxy)-phosphoryl)-L-alaninat (82**)**



Das geschützte 5-Aza-carbaC **80** (69 mg, 155 μ mol, 1.0 eq., als 1.6 eq. NEt_3 -Addukt) wurde in 2.40 mL trockenem THF suspendiert und auf -78°C gekühlt. Anschließend wurden 855 μ L tBuMgCl -Lösung (1.0 M in THF, 855 μ mol, 5.5 eq.) zugetropft, das Gemisch direkt auf 0°C erwärmt und 30 min bei dieser Temperatur verrührt. Daraufhin wurde das Gemisch wieder auf -78°C gekühlt und tropfenweise mit 538 μ L einer Lösung des Phosphoroxarylchlorids^[138] **81** (1.0 M in THF, 538 μ mol, 3.5 eq.) versetzt. Das Reaktionsgemisch wurde langsam auf Raumtemperatur erwärmen gelassen und anschließend 17 h bei Raumtemperatur verrührt. Die Reaktion wurde durch Zugabe von H_2O (20 mL) gestoppt und die wässrige Phase wurde mit EtOAc (3 x 25 mL) extrahiert. Die vereinigten organischen Phasen wurden mit gesättigter wässriger Natriumchloridlösung (25 mL) gewaschen, über wasserfreiem Na_2SO_4 getrocknet und schließlich *in vacuo* zur Trockene eingeeengt. Der Rückstand wurde mittels Säulenchromatographie an Silicagel aufgereinigt ($\text{CH}_2\text{Cl}_2:\text{MeOH}$ 60:1 \rightarrow $\text{CH}_2\text{Cl}_2:\text{MeOH}$ 50:1 \rightarrow $\text{CH}_2\text{Cl}_2:\text{MeOH}$ 40:1), um 83 mg des Reaktionsprodukts **82** (78 μ mol, 51%, Addukt mit 5.0 eq. NEt_3) als Gemisch der beiden Diastereomere am P in Form eines farblosen Feststoffs zu erhalten.

^1H -NMR (800 MHz, CDCl_3): δ/ppm = 8.05 (2 x s, 1H), 7.32 – 7.27 (m, 2H, OPh), 7.23 – 7.17 (m, 2H, OPh), 7.15 – 7.07 (m, 1H, OPh), 5.04 – 4.87 (m, 2H, 2'-H, 4'''-H), 4.59 – 4.54 (m, 1H, 3'-H), 4.31 – 4.13 (m, 3H, 1'-H, 5'-H), 4.02 – 3.93 (m, 1H, 2'''-H), 3.78 (m, 1H, NH), 3.09 (q, J = 7.4 Hz, 30H, $\text{N}(\text{CH}_2\text{CH}_3)_3$), 2.53 – 2.17 (m, 3H, 4'-H, 6'-H), 1.49 – 1.47 (m, 3H, 1''-H), 1.39 (t, J = 7.3 Hz, 45H, $\text{N}(\text{CH}_2\text{CH}_3)_3$), 1.37 – 1.34 (m, 3H, 3''-H), 1.26 – 1.24 (m, 3H, 1''-H), 1.21 – 1.18 (m, 6H, 5'''-H).

^{13}C -NMR (201 MHz, CDCl_3): δ/ppm = 173.10 ($1''''\text{-C}$), 165.39 (4-C), 158.57 (6-C), 153.13 (O- C_{Ar}), 150.83 (2-C), 129.73 (CH_{Ar}), 124.99 (CH_{Ar}), 120.45 (CH_{Ar}), 113.57 ($2''\text{-C}$), 82.42 ($2'\text{-C}$), 81.17 ($3'\text{-C}$), 69.28 ($4'''\text{-C}$), 67.50 ($1'\text{-C}$), 65.93 ($5'\text{-C}$), 50.50 ($2'''\text{-C}$), 45.91 ($\text{N}(\text{CH}_2\text{CH}_3)_3$), 44.93 ($4'\text{-C}$), 32.34 ($6'\text{-C}$), 29.79 ($5'''\text{-C}$), 27.73 ($1''\text{-C}$), 25.36 ($1'''\text{-C}$), 21.72 ($5'''\text{-C}$), 21.13 ($3'''\text{-C}$), 8.75 ($\text{N}(\text{CH}_2\text{CH}_3)_3$).

^{31}P -NMR (162 MHz, CDCl_3): δ/ppm = 2.46, 2.36.

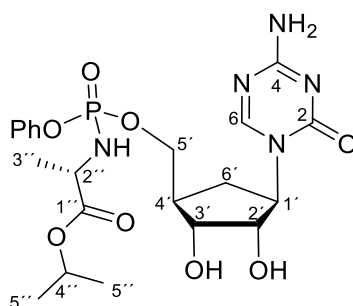
HRMS (ESI $^+$): m/z berechnet für $\text{C}_{24}\text{H}_{35}\text{N}_5\text{O}_8\text{P}^+$ 552.2218 $[\text{M}+\text{H}]^+$; gefunden: 552.2211.

HRMS (ESI $^-$): m/z berechnet für $\text{C}_{24}\text{H}_{33}\text{N}_5\text{O}_8\text{P}^-$ 550.2072 $[\text{M}-\text{H}]^-$; gefunden: 550.2074.

IR (ATR): ν (cm^{-1}) = 3379 (m), 2980 (s), 2946 (m), 2606 (s), 2498 (s), 1640 (s), 1476 (s), 1398 (m), 1383 (m), 1212 (s), 1158 (m), 1107 (w), 1070 (m), 1037 (s), 929 (m), 802 (w), 730 (m).

R_f (CH_2Cl_2 :MeOH 20:1): 0.25.

Isopropyl((((1R,2R,3S,4R)-4-(4-amino-2-oxo-1,3,5-triazin-1(2H)-yl)-2,3-dihydroxycyclopentyl)methoxy)(phenoxy)phosphoryl)-L-alaninat (83)



Zu der Isopropylidenacetal-geschützten Verbindung **82** (38 mg, 69 μmol , 1.0 eq.) wurden 1.0 mL verdünnte TFA (80% in ddH_2O) zugetropft und das Gemisch wurde 20 min bei 30 °C verrührt. Anschließend wurde das Gemisch in 5.0 mL ddH_2O gegossen und gefriergetrocknet. Der erhaltene Rückstand wurde in einem Gemisch aus 500 μL MeCN, 500 μL DMSO und 4.0 mL ddH_2O (enthaltend 0.1% TFA) gelöst, filtriert und mittels *reversed-phase* HPLC (5% MeCN in ddH_2O mit 0.1% TFA auf 60% MeCN in ddH_2O mit 0.1% TFA in 45 min, Fluss 5 mL/min) aufgereinigt, um 10.4 mg der Zielverbindung **83** (20 μmol , 29 %) in Form eines farblosen Feststoffs zu erhalten.

¹H-NMR (400 MHz, D₃CCN): δ/ppm = 8.11 (s, 1H), 7.50 (s, 1H, NH₂), 7.41 – 7.29 (m, 2H, OPh), 7.25 – 7.15 (m, 3H, OPh), 6.61 (s, 1H, NH₂), 4.98 – 4.88 (m, 1H, 4''-H), 4.39 – 4.23 (m, 3H, 1'-H, 2'-H, NH), 4.20 – 4.07 (m, 2H, 5'-H), 3.97 (dt, *J* = 5.2, 3.8 Hz, 1H, 3'-H), 3.93 – 3.82 (m, 1H, 2''-H), 2.33 – 2.09 (m, 2H, 4'-H, 6'-H), 1.83 – 1.67 (m, 1H, 6'-H), 1.30 – 1.26 (m, 3H, 3''-H), 1.19 (dt, *J* = 6.2, 3.4 Hz, 6H, 5''-H).

¹³C-NMR (101 MHz, D₃CCN): δ/ppm = 174.00 (1''-C), 165.92 (4-C), 159.92 (6-C), 153.84 (2-C), 151.97 (O-C_{Ar}), 130.62 (CH_{Ar}), 125.75 (CH_{Ar}), 121.36 (CH_{Ar}), 121.31 (CH_{Ar}), 74.12 (2'-C), 72.50 (3'-C), 69.64 (4''-C), 68.53 (5'-C), 64.96 (1'-C), 51.44 (2''-C), 44.44 (4'-C), 28.08 (6'-C), 21.89 (5''-C), 21.81 (5''-C), 20.80 (3''-C).

³¹P-NMR (162 MHz, D₃CCN): δ/ppm = 3.08, 2.92.

HRMS (ESI⁺): *m/z* berechnet für C₂₁H₃₁N₅O₈P⁺ 512.1905 [M+H]⁺; gefunden: 512.1901.

HRMS (ESI⁻): *m/z* berechnet für C₂₁H₂₉N₅O₈P⁻ 510.1759 [M-H]⁻; gefunden: 510.1757.

Zur Trennung der diastereomeren Verbindungen mittels semipräparativer HPLC wurde eine Probe des Rohprodukts von **83** in einem Gemisch aus 10% MeCN, 10% DMSO und 80% ddH₂O gelöst und bei einem Fluss von 5.0 mL/min mit einem Gradienten eines 0.01% TFA enthaltenden Eluentensystems von 5% MeCN in ddH₂O auf 60% MeCN in ddH₂O innerhalb von 45 min aufgereinigt.

Untersuchungen zur Hydrolysestabilität von cAzaC **79**

Zur Untersuchung der Hydrolysestabilität von cAzaC **79** sowie der kommerziell erhältlichen und bereits bekannten Verbindung 5-Aza-C wurden beide Verbindungen jeweils in 100 mM wässriger Lösung (pH = 7) bei Raumtemperatur inkubiert. Zur Untersuchung wurden jeweils 3 µL dieser Lösungen zu bestimmten Zeitpunkten (0 h, 1,5 h, 3 h, 4,5 h, 6 h, 24 h, 6 d) entnommen, mit 20 µL ddH₂O verdünnt und anschließend wurden 18 µL der verdünnten Lösungen mittels analytischer HPLC untersucht. Dabei wurde bei einer Flussrate von 0.5 mL/min ein Gradient von 0% MeCN auf 5% MeCN in ddH₂O innerhalb von 25 min eingesetzt und das Produktgemisch bei den Wellenlängen 248 nm und 260 nm analysiert.

6.6. Untersuchung der Basenpaarungseigenschaften von Inosin mittels eines isotonenmarkierten Inosin-RNA-Duplexes

6.6.1. Synthese eines ^{15}N -isotonenmarkierten Inosin-Phosphoramidits, Oligonukleotidsynthese und erste NMR-spektroskopische Ergebnisse

Trotz des häufigen Auftretens der Inosin-Modifikation, insbesondere in doppelsträngiger mRNA, sind dessen Basenpaarungseigenschaften im Doppelstrang bisher nicht ausreichend untersucht. Es ist bekannt, dass Inosin bevorzugt mit einem C im Gegenstrang die Basenpaarung ausbildet,^[98] jedoch entsteht beim *A-zu-I-editing* der doppelsträngigen mRNA eine Basenfehlpaarung, bei der Inosin gegenüber einem Uridin positioniert ist.^[95] Im Rahmen dieses Projekts sollte diese I-U-Basenpaarung durch NMR-spektroskopische Messungen eines Inosin-enthaltenden RNA-Duplexes in Kooperation mit der Arbeitsgruppe von *Prof. Michael Sattler* an der *TU München* genauer untersucht werden. Dabei sollte ermittelt werden, ob die *N1*-Position des Inosins an dieser Basenpaarung direkt beteiligt ist.

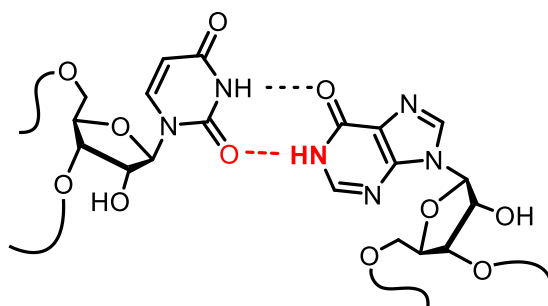


Abbildung 6-39: Schema der Untersuchung des Basenpaarungsverhaltens des U-I-Basenpaares ausgehend vom *N1*-Atoms am Inosin.

Dazu sollte zunächst ein an der *N1*-Position isotonenmarkiertes ^{15}N -Inosin-Phosphoramidit synthetisiert werden, das anschließend zur RNA-Oligonukleotidsynthese verwendet werden konnte. Nach der Synthese von zwei komplementären RNA-Strängen, bei denen isotonenmarkierte Inosin-Modifikationen gegenüber von Uridinbasen positioniert werden sollten, sollten die Oligonukleotide dann im Doppelstrang im Hinblick auf die Basenpaarungseigenschaften des Inosins NMR-spektroskopisch untersucht werden.

Bevor das Phosphoramidit für die RNA-Festphasensynthese hergestellt werden konnte, musste zunächst die Synthese des gezielt an der *N1*-Position ^{15}N -isotonenmarkierten Inosin-Phosphoramidits entwickelt werden. Ausgehend von Imidazol-4-carbonsäure wurde anhand bereits bekannter Literaturvorschriften $^{15}\text{N1}$ -

Hypoxanthin in vier Stufen synthetisiert (siehe Abbildung 6-40).^[141,142] Hierzu wurde Imidazol-4-carbonsäure **84** in Gegenwart von Ammoniumnitrat in konzentrierter Schwefelsäure selektiv an der 5-Position nitriert. Die Nitrogruppe sollte dabei als maskierte Aminofunktion dienen, die im späteren Verlauf der Synthese durch Reduktion freigelegt werden sollte. Zuvor wurde allerdings die ¹⁵N-Isotopenmarkierung eingeführt. Dazu wurde zuerst die Carbonsäureeinheit von **85** durch Reaktion mit Carbonyldiimidazol aktiviert und anschließend durch Umsetzung mit isotoopenmarkiertem ¹⁵N-Ammoniumchlorid (¹⁵NH₄Cl) zum Carbonsäureamid **86** umgewandelt. Daraufhin folgte die Reduktion der Nitrogruppe an 5-Position zum Amin durch Wasserstoff in Gegenwart von Palladium auf Kohle und abschließend der Ringschluss zum Hypoxanthin **88**. Der Ringschluss wurde dabei in einer zweifachen Kondensationsreaktion von **87** mit Diethoxymethylacetat (DEMA) in Gegenwart von Ameisensäure bei erhöhter Temperatur durchgeführt. Insgesamt konnte das isotoopenmarkierte Hypoxanthin **88** nur in einer geringen Gesamtausbeute von 9% ausgehend von Imidazol-4-carbonsäure **84** in einem vierstufigen Verfahren erhalten werden.

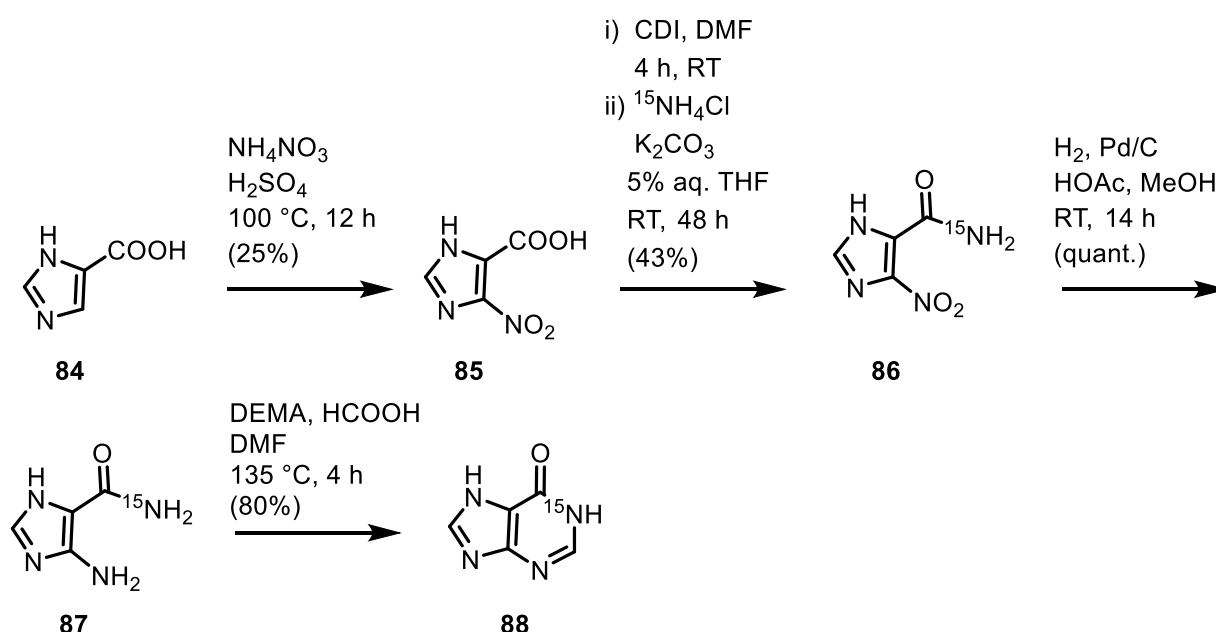


Abbildung 6-40: Synthese von ¹⁵N₁-Hypoxanthin **88**.

Insbesondere die geringe Ausbeute der Nitrierung im ersten synthetischen Schritt senkt dabei die Gesamtausbeute, da das Nitrierungsprodukt **85** in nur 25% erhalten werden konnte. Die starke Polarität und gleichzeitig schlechte Löslichkeit der

einzelnen Verbindungen, die nur Aufreinigungen mittels Umkristallisation zuließen, limitierten ebenfalls Spielraum für mögliche Optimierungen. Nichtsdestotrotz konnte diese vierstufige Purinsynthese von **88** in großem Maßstab ausgehend von dem kommerziell erhältlichen Startmaterial **84** durchgeführt werden, um ausreichend Material für die restlichen synthetischen Transformationen zu erhalten.

Im nächsten Schritt wurde das Hypoxanthin unter Verwendung kommerziell erhältlicher 1-O-Acetyl-2,3,5-tri-O-benzoyl-β-D-ribofuranose **89** als Glykosyldonor unter leicht modifizierten Standardbedingungen zum geschützten Inosin-Derivat nukleosidiert. Dabei wurde die Reaktion bei erhöhter Temperatur im Lösungsmittel Toluol unter Verwendung von Bis(trimethylsilyl)acetamid und TMSOTf durchgeführt.

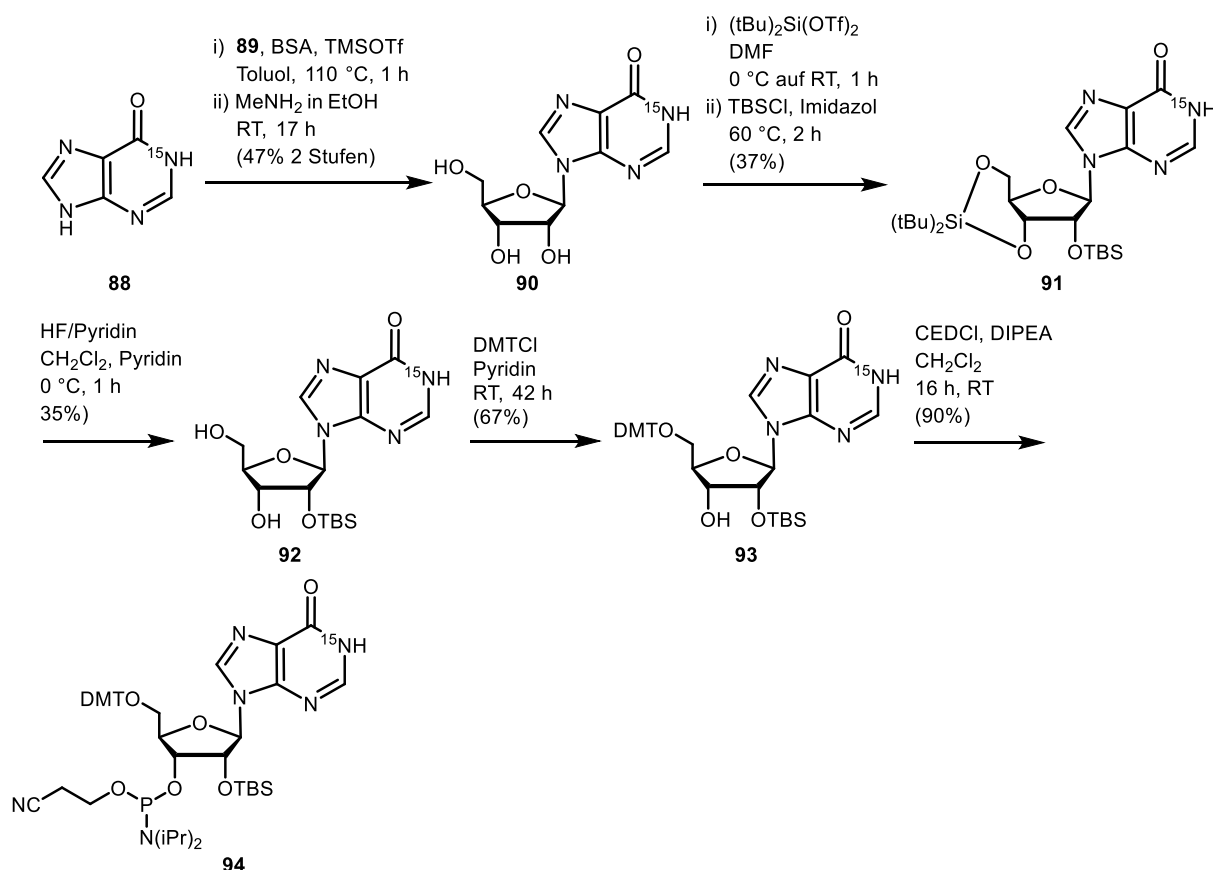


Abbildung 6-41: Syntheschema des ¹⁵N1-Inosin-Phosphoramidits **94**.

Die anschließende Entschützung von den Benzoyl-Gruppen des erhaltenen Nukleosid-Derivats durch ethanolische Methylaminlösung zu **90** und direkt nachfolgende Silyl-Schützung der freigelegten Hydroxygruppen lieferte das geschützte Inosin-Derivat **91**. Die Silyl-Schützung konnte dabei analog zur Literatur im Eintopfverfahren durchgeführt werden,^[95] wobei zunächst bei 0 °C die

Silylklammer an 3'- und 5'-Position und anschließend bei erhöhter Temperatur die TBS-Gruppe an 2'-Position eingeführt wurden. Der Einsatz der 3',5'-(*O*-di-*tert*-butylsilylandiyl)-Schutzgruppe erlaubte die selektive Freilegung der 3'- und 5'-Hydroxygruppen unter Verwendung von Fluorwasserstoff in Pyridin bei 0 °C und einer kurzen Reaktionsdauer von einer Stunde. Im Anschluss wurde die DMT-Schutzgruppe selektiv an der Hydroxygruppe in 5'-Position unter Standardbedingungen eingeführt. Erwähnenswert ist, dass zur säulenchromatographischen Aufreinigung des erhaltenen Rohprodukts **93** ein Eluentensystem aus Isohexan, Aceton und Dichlormethan eingesetzt werden musste.^[95] Die Verwendung eines Eluentensystems aus Dichlormethan und Methanol unter Verwendung von Triethylamin zur Deaktivierung des Silicagels hingegen führte zu einer partiellen Migration der TBS-Schutzgruppe von der 2'-Position zur 3'-Position. Um eine Auftrennung dieses Gemisches zu umgehen, wurde das oben beschriebene Laufmittelsystem verwendet. Abschließend erfolgte die Synthese des Phosphoramidits **94** unter Standardbedingungen (siehe Abbildung 6-41). Allerdings konnte der Umsatz der Reaktion dünnschichtchromatographisch auf Grund der Kopolarität von Startmaterial **93** und Reaktionsprodukt **94** nicht verfolgt werden. Des Weiteren wurde das Reaktionsprodukt **94** als Addukt mit dem zur Säulenchromatographie der Reaktionsmischung verwendeten Triethylamin (1.7 eq.) erhalten. Diese Adduktbildung wurde schon bei der Synthese eines Inosin-Phosphoramidits mit anderem Isotopenmarkierungsmuster von *Dallmann et al.* beobachtet.^[95] Nichtsdestotrotz konnte das erhaltene Phosphoramidit **94** zur RNA-Festphasensynthese verwendet werden.

Die Oligonukleotid-Festphasensynthese wurde für das Inosin-Phosphoramidit **94** leicht von den Standardbedingungen abgewandelt. Phosphoramidit **94** wurde auf Grund seiner Löslichkeitseigenschaften nicht als Lösung in Acetonitril sondern als Lösung in Dichlormethan eingesetzt. Bei der Festphasensynthese wurden des Weiteren auch verdoppelte Kopplungszeiten und die doppelte Stoffmenge des Phosphoramidits zur Kopplung verwendet (siehe Kapitel 6.6.3). Unter diesen modifizierten Synthesebedingungen konnten die erforderlichen 20mer-Oligonukleotide synthetisiert werden, die den zu untersuchenden RNA-Doppelstrang ausbilden sollten. Eines dieser Oligonukleotide enthielt dabei eine Inosin-Modifikation, wohingegen der Gegenstrang drei Inosin-Basen aufwies. Die Sequenzen der Oligonukleotide basieren auf einem Fragment doppelsträngiger RNA,

die in den Exons 2 und 3 des α -Tropomyosins von Ratten gefunden werden kann.^[95] Bei diesem RNA-Motiv ist bekannt, dass *A-zu-I editing* in diesem Sequenzbereich innerhalb von vier Basen viermal vorkommen kann.^[95] In der Folge können innerhalb von vier Basenpaaren alle vier Inosin-Modifikationen gegenüber von Uridin-Basen positioniert sein. Dieser Sequenzkontext sollte für die NMR-spektroskopischen Untersuchungen des Basenpaarungsverhaltens der *N1*-Position von Inosin eingesetzt werden. Die NMR-Messungen der isotopenmarkierten Oligonukleotide wurden dabei in Kooperation mit der Arbeitsgruppe von *Prof. Michael Sattler* an der *TU München* durchgeführt.

Erste NMR-spektroskopische Untersuchungen in der Arbeitsgruppe von *Prof. Sattler* konnten dabei die erfolgreiche Synthese des RNA-Duplexes bestätigen. Abbildung 6-42 zeigt ein dort gemessenes NOESY-Spektrum der Imino-Region der untersuchten RNA (blau) gemessen an einem 600 MHz-Spektrometer. Die roten Signale wurden bei Messungen von sequenzgleicher RNA, die ein unterschiedliches Isotopenmarkierungsmuster aufwies, in der Arbeitsgruppe von *Prof. Sattler* gemessen. Auf Grund des exakten Überlapps der Signale der beiden RNA-Proben konnte die Struktur der hier synthetisierten RNA im Vergleich zur Kontrollprobe mit identischer Sequenz bestätigt werden. Durch die Verwendung eines leicht sauren Phosphatpuffers (pH 6.4) konnten im Rahmen dieser Messungen vergleichsweise sogar noch mehr Iminoprotonen als bei den früheren Messungen bei neutralem pH gefunden werden.

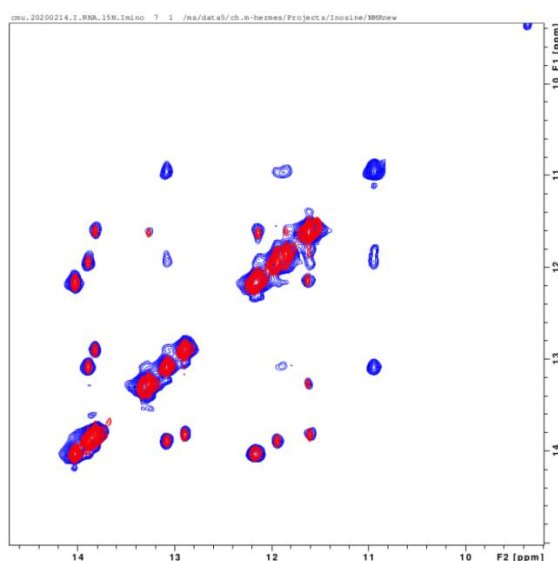


Abbildung 6-42: NOESY-Spektrum der Imino-Region der untersuchten RNA (blau) und der Vergleichsprobe aus *Dallmann et al.*^[95].

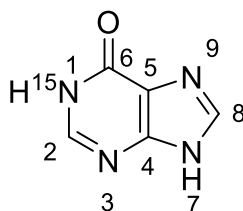
Einige der hier neu entdeckten Signale konnten Raumwechselwirkungen zwischen einer der Inosin-Modifikationen mit Iminoprotonen der Basen des Gegenstrangs zugeordnet werden. Nichtsdestotrotz konnten bisher nicht alle Iminoprotonen der inkorporierten Inosin-Basen detektiert werden, was möglicherweise durch überlappende Resonanzen der Iminoprotonen der Inosinbasen bedingt sein kann. Des Weiteren konnten bisher die ^{15}N -Isotopenmarkierungen an den *N1*-Positionen der Inosin-Modifikationen über ^{15}N -2D-NMR-Experimente nicht detektiert werden. Im Zuge dessen werden gegenwärtig weitere Studien und Messungen in der Arbeitsgruppe von *Prof. Sattler* an der *TU München* durchgeführt. Ziel dabei ist zunächst die ^{15}N -Resonanzen zu detektieren und anschließend basierend darauf dynamische Effekte dieser Signale NMR-spektroskopisch zu untersuchen. Damit könnten möglicherweise Aussagen getroffen werden, ob die *N1*-Position der Inosin-Modifikation bei der Basenpaarung zwischen Inosin und Uridin direkt eine Wasserstoffbrückenbindung ausbildet.

6.6.2. Projektbeiträge

Die Syntheseroute zum ^{15}N -isotopenmarkierten Inosin-Phosphoramidit **94** wurde von mir entwickelt und ebenso wie die anschließende Oligonukleotidsynthese von mir in Zusammenarbeit mit *Florian Schelter* umgesetzt. NMR-spektroskopische Untersuchungen des RNA-Duplexes werden in der Arbeitsgruppe von *Prof. Michael Sattler* an der *TU München* durchgeführt.

6.6.3. Experimentalteil

[¹⁵N1]-Hypoxanthin (**88**)

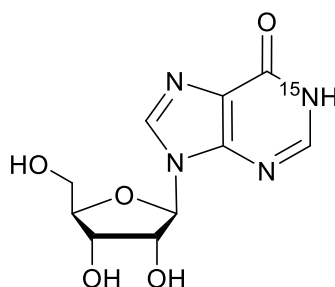


[¹⁵N1]-Hypoxanthin **88** wurde ausgehend von Imidazol-4-carbonsäure **84** in einem vierstufigen Verfahren gemäß bestehender Literaturvorschriften synthetisiert.^[141,142]

¹H-NMR (400 MHz, D₂O/NaOD): δ/ppm = 7.77 – 7.72 (m, 1H, 2-H), 7.54 (s, 1H, 8-H).

Weitere analytische Daten der Verbindung **88** sind in der Literatur zu finden.^[142]

[¹⁵N1]-Inosin (**90**)



In einem ausgeheizten Kolben wurden [¹⁵N1]-Hypoxanthin **88** (0.515 g, 3.76 mmol, 1.0 eq.) und 1-O-Acetyl-2,3,5-tri-O-benzoyl-β-D-ribofuranose **89** (2.09 g, 4.14 mmol, 1.1 eq.) in 30 mL trockenem Toluol gelöst. Das Gemisch wurde mit BSA (5.60 mL, 22.6 mmol, 6.0 eq.) versetzt 30 min bei 110 °C refluxiert. Daraufhin wurde das Gemisch auf Raumtemperatur abkühlen gelassen und mit TMSOTf (2.10 mL, 11.3 mmol, 3.0 eq.) versetzt. Anschließend wurde wieder für 30 min auf 110 °C erwärmt, nach Abkühlen auf Raumtemperatur wurde das Lösungsmittel *in vacuo* entfernt, der Rückstand wurde in CH₂Cl₂ (150 mL) aufgenommen und mit gesättigter wässriger Na₂CO₃-Lösung (150 mL) gewaschen. Die wässrige Phase wurde weiterhin mit CH₂Cl₂ (2 x 150 mL) extrahiert, die vereinigten organischen Phasen wurden über Na₂SO₄ getrocknet und Lösungsmittel wurden *in vacuo* entfernt. Der erhaltene Rückstand wurde durch Silicagel (CH₂Cl₂:MeOH 60:1 → CH₂Cl₂:MeOH 20:1) gefiltert, um als Rohprodukt [¹⁵N1]-Inosin-2',3',5'-tribenzoat (1.02 g, 1.75 mmol) als gelben Feststoff zu liefern.

Das geschützte Inosin (1.02 g, 1.75 mmol) wurde in 15.0 mL ethanolischer Methylaminlösung (33 Gew.-%) gelöst und 17 h bei Raumtemperatur verrührt. Anschließend wurde das Reaktionsgemisch *in vacuo* zur Trockene eingeeengt, der Rückstand in ddH₂O (30 mL) und CH₂Cl₂ (30 mL) aufgenommen, die Phasen getrennt und die wässrige Phase mit CH₂Cl₂ gewaschen (4 x 30 mL). Anschließend wurde die wässrige Phase gefriergetrocknet, um das vollständig entschützte ¹⁵N-Inosin **90** (471 mg, 1.75 mmol, 47% über 2 Stufen) als gelblichen Feststoff zu liefern, der ohne weitere Aufreinigung für die nächste Stufe verwendet werden konnte.

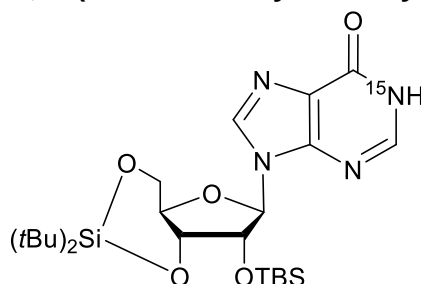
¹H-NMR (400 MHz, D₂O): δ/ppm = 8.31 (s, 1H, 8-H), 8.22 – 8.17 (m, 1H, 2-H), 6.06 (d, *J* = 5.7 Hz, 1H, 1'-H), 4.73 (t, *J* = 5.5 Hz, 1H, 2'-H), 4.40 (dd, *J* = 5.2, 3.9 Hz, 1H, 3'-H), 4.24 (td, *J* = 3.9, 2.9 Hz, 1H, 4'-H), 3.92 – 3.77 (m, 2H, 5'-H).

¹³C-NMR (101 MHz, D₂O): δ/ppm = 148.43 (6-C), 146.09 – 145.99 (m, 2-C), 140.15 (8-C), 128.67(4/5-C), 127.31 (4/5-C), 88.23 (1'-C), 85.49 (4'-C), 73.92 (2'-C), 70.27 (3'-C), 61.18 (5'-C).

HRMS (ESI⁺): *m/z* berechnet für C₁₀H₁₂N₃¹⁵NO₅Na⁺ [M+Na]⁺ 292.0670; gefunden: 292.0672.

HRMS (ESI⁻): *m/z* berechnet für C₁₀H₁₁N₃¹⁵NO₅⁻ [M-H]⁻ 268.0705; gefunden: 268.0705.

2'-(*O*-*tert*-Butyldimethylsilyl)-3',5'-(*O*-di-*tert*-butylsilandiyl)-[¹⁵M]-inosin (91**)**



Das entschützte, isotonenmarkierte Inosin **90** (476 mg, 1.76 mmol, 1.0 eq.) wurde in 9.0 mL trockenem DMF gelöst und auf 0 °C gekühlt. Daraufhin wurde (tBu)₂Si(OTf)₂ (691 µL, 2.12 mmol, 1.2 eq.) zugetropft und das Reaktionsgemisch 45 min bei 0 °C verrührt. Anschließend wurde Imidazol (547 mg, 8.03 mmol, 5.0 eq.) zugegeben und das Gemisch wurde über 30 min auf Raumtemperatur erwärmt. Schließlich wurde dem Gemisch TBS-Cl (748 mg, 3.18 mmol, 1.8 eq.) zugesetzt und für 2 h auf 60 °C erwärmt. Nach vollständigem Umsatz des Startmaterials wurde das Gemisch in H₂O

(100 mL) gegossen und mit EtOAc (3 x 100 mL) extrahiert. Die vereinigten organischen Phasen wurden mit gesättigter wässriger Natriumchloridlösung (100 mL) gewaschen, über Na₂SO₄ getrocknet und *in vacuo* zur Trockene eingengt. Eine Aufreinigung des Rückstandes durch Säulenchromatographie an Silicagel (CH₂Cl₂ → CH₂Cl₂:MeOH 80:1 → CH₂Cl₂:MeOH 40:1 → CH₂Cl₂:MeOH 20:1) lieferte das geschützte Reaktionsprodukt **91** (340 mg, 0,65 mmol, 37%) als gelblichen Schaum.

¹H-NMR (800 MHz, DMSO-*d*₆): δ/ppm = 12.52 – 12.33 (m, 1H, NH), 8.29 (s, 1H, 8-H), 8.07 – 8.04 (m, 1H, 2-H), 5.95 – 5.89 (m, 1H, 1'-H), 4.61 (d, *J* = 5.1 Hz, 1H, 2'-H), 4.54 – 4.48 (m, 1H, 3'-H), 4.37 (dd, *J* = 8.5, 4.4 Hz, 1H, 5'-H), 4.06 – 3.97 (m, 2H, 4'-H, 5'-H), 1.06 (s, 9H, Si(C(CH₃)₃)₂), 1.01 (s, 9H, Si(C(CH₃)₃)₂), 0.87 (s, 9H, Si(CH₃)₂C(CH₃)₃), 0.10 (s, 3H, Si(CH₃)₂C(CH₃)₃), 0.08 (s, 3H, Si(CH₃)₂C(CH₃)₃).

¹³C-NMR (201 MHz, DMSO-*d*₆): δ/ppm = 146.44 (2-C), 139.26 (8-C), 91.50 (1'-C), 75.76 (3'-C), 75.33 (2'-C), 74.69 (4'-C), 67.37 (5'-C), 27.79 (Si(C(CH₃)₃)₂), 27.33 (Si(C(CH₃)₃)₂), 26.13 (Si(CH₃)₂C(CH₃)₃), -4.41 (Si(CH₃)₂C(CH₃)₃), -4.73 (Si(CH₃)₂C(CH₃)₃).

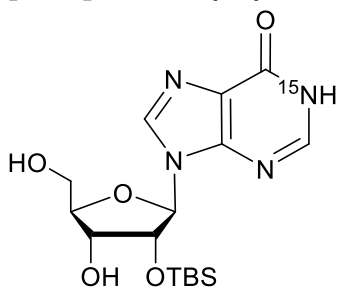
HRMS (ESI⁺): *m/z* berechnet für C₂₄H₄₃N₃¹⁵NO₅Si₂⁺ [M+H]⁺ 524.2737; gefunden: 524.2736.

HRMS (ESI⁻): *m/z* berechnet für C₂₄H₄₁N₃¹⁵NO₅Si₂⁻ [M-H]⁻ 522.2591; gefunden: 522.2598.

IR (ATR): *v* (cm⁻¹) = 29633 (s), 2891 (w), 2859 (s), 2358 (w), 1694 (s), 1566 (w), 1472 (m), 1364 (m), 1257 (m), 1206 (w), 1167 (m), 1142 (m), 1059 (s), 1000 (m), 939 (w), 891 (m), 826 (s), 781 (m), 752 8w), 733 (m), 681 (w).

R_f (CH₂Cl₂:MeOH 15:1): 0.59.

2'-(*O*-*tert*-Butyldimethylsilyl)-[¹⁵M]-inosin (**92**)



Das vollständig geschützte Inosin-Derivat **91** (316 mg, 0.60 mmol, 1.0 eq.) wurde in 15.0 mL trockenem CH₂Cl₂ und 6.0 mL trockenem Pyridin gelöst und daraufhin bei 0 °C mit Fluorwasserstoff in Pyridin (70%, 217 µL, 9.05 mmol, 15.0 eq.) versetzt. Das Reaktionsgemisch wurde 1 h bei 0 °C verrührt und anschließend wurde die Reaktion durch Zugabe von TMSOMe (2.66 mL, 19.31 mmol, 32.0 eq.) gestoppt. Nach Einengen des Gemisches *in vacuo* wurde der Rückstand über Säulenchromatographie an Silicagel (CH₂Cl₂:MeOH 60:1 → CH₂Cl₂:MeOH 20:1) aufgereinigt, um das an 3'- und 5'-Position entschlützte Nukleosid **92** (83 mg, 0.21 mmol, 35%) als farblosen Feststoff zu liefern.

¹H-NMR (400 MHz, CD₃OD): δ/ppm = 8.37 (s, 1H, 8-H), 8.14 – 8.03 (m, 1H, 2-H), 5.99 (d, *J* = 5.9 Hz, 1H, 1'-H), 4.72 (t, *J* = 5.4 Hz, 1H, 2'-H), 4.24 (dd, *J* = 5.0, 3.1 Hz, 1H, 3'-H), 4.14 (q, *J* = 2.9 Hz, 1H, 4'-H), 3.86 (dd, *J* = 12.4, 2.8 Hz, 1H, 5'-H), 3.74 (dd, *J* = 12.4, 2.9 Hz, 1H, 5'-H), 0.79 (s, 9H, SiC(CH₃)₃), -0.05 (s, 3H, SiCH₃), -0.16 (s, 3H, SiCH₃).

¹³C-NMR (101 MHz, CD₃OD): δ/ppm = 158.80 – 158.69 (m, 6-C), 149.60 (4-C), 146.97 – 146.89 (m, 2-C), 141.09 (8-C), 126.12 – 126.04 (m, 5-C), 90.58 (1'-C), 87.91 (4'-C), 77.85 (2'-C), 72.45 (3'-C), 62.99 (5'-C), 26.13 (SiC(CH₃)₃), 18.96 (SiC(CH₃)₃), -4.84 (SiCH₃), -5.19 (SiCH₃).

¹⁵N-NMR (41 MHz, CD₃OD): δ/ppm = -280.58.

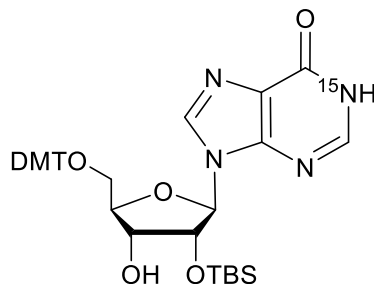
HRMS (ESI⁺): *m/z* berechnet für C₁₆H₂₇N₃¹⁵NO₅Si⁺ [M+H]⁺ 384.1716; gefunden: 384.1714.

HRMS (ESI⁻): *m/z* berechnet für C₁₆H₂₅N₃¹⁵NO₅Si⁻ [M-H]⁻ 382.1570; gefunden: 382.1573.

IR (ATR): ν (cm⁻¹) = 3115 (w), 2927 (m), 2856 (m), 2360 (w), 2250 (w), 1691 (s), 1588 (w), 1545 (m), 1507 (w), 1472 (w), 1421 (w), 1374 (w), 1321 (w), 1253 (m), 1216 (m), 1150 (m), 1087 (s), 1029 (w), 1005 (w), 867 (m), 837 (s), 787 (s), 715 (w), 692 (w).

R_f (CH₂Cl₂:MeOH 15:1): 0.25.

5'-O-(4,4'-Dimethoxytrityl)-2'-(O-*tert*-butyldimethylsilyl)-[¹⁵N]-inosin (93**)**



Unter Ar-Atmosphäre wurde das an 2'-Position TBS-geschützte Inosin Derivat **92** (83 mg, 0.21 mmol, 1.0 eq.) in 2.2 mL trockenem Pyridin gelöst und auf 0 °C gekühlt. Anschließend wurde 4,4'-Dimethoxytritylchlorid (DMT-Cl, 81 mg, 0.24 mmol, 1.1 eq.) zugegeben und das Reaktionsgemisch wurde auf Raumtemperatur erwärmt und 18 h verrührt. Daraufhin wurde erneut DMT-Cl (123 mg, 0.36 mmol, 1.75 eq.) zugegeben und für weitere 24 h bei Raumtemperatur verrührt. Nach vollständigem Umsatz des Startmaterials wurde die Reaktionsmischung auf gesättigte wässrige NaHCO₃-Lösung (50 mL) gegossen und mit CH₂Cl₂ (3 x 50 mL) extrahiert. Die vereinigten organischen Phasen wurden über Na₂SO₄ getrocknet, mit Toluol (30 mL) versetzt und *in vacuo* zur Trockene eingengt. Der Rückstand wurde mittels Säulenchromatographie an Silicagel (CH₂Cl₂ → iHex:Aceton:CH₂Cl₂ 35:40:25) aufgereinigt, um die DMT-geschützte Verbindung **93** (96 mg, 0.14 mmol, 67%) als farblosen Feststoff zu liefern.

¹H-NMR (400 MHz, CD₂Cl₂): δ/ppm = 11.67 (s, 1H, NH), 7.99 – 7.94 (2 s, 2H, 2-H, 8-H), 7.49 – 7.43 (m, 2H, C-H_{Ar}(DMT)), 7.37 – 7.20 (m, 7H, C-H_{Ar}(DMT)), 6.86 – 6.80 (m, 4H, C-H_{Ar}(DMT)), 5.98 (d, *J* = 5.3 Hz, 1H, 1'-H), 4.87 (t, *J* = 5.2 Hz, 1H, 2'-H), 4.32 (q, *J* = 4.3 Hz, 1H, 3'-H), 4.24 (q, *J* = 3.7 Hz, 1H, 4'-H), 3.78 (s, 6H, 2 x OCH₃), 3.46 (dd, *J* = 10.6, 3.2 Hz, 1H, 5'-H), 3.39 (dd, *J* = 10.7, 4.3 Hz, 1H, 5'-H), 0.85 (s, 9H, SiC(CH₃)₃), 0.02 (s, 3H, SiCH₃), -0.10 (s, 3H, SiCH₃).

¹³C-NMR (101 MHz, CD₂Cl₂): δ/ppm = 159.25 (C_{Ar}(DMT)), 158.75 (6-C), 149.27 (4-C), 145.35 – 145.08 (m, 2-C), 139.36 (8-C), 136.12 (C_{Ar}(DMT)), 130.62 (C_{Ar}(DMT)), 128.60 (C_{Ar}(DMT)), 128.44 (C_{Ar}(DMT)), 127.45 (C_{Ar}(DMT)), 125.82 (5-C), 113.67 (C_{Ar}(DMT)), 88.98 (1'-C), 87.14 (C_{Ar3}), 84.75 (4'-C), 76.55 (2'-C), 72.03 (3'-C), 64.06 (5'-C), 55.77 (OCH₃), 25.87 (SiC(CH₃)₃), 18.34(SiC(CH₃)₃), -4.70 (SiCH₃), -4.87 (SiCH₃).

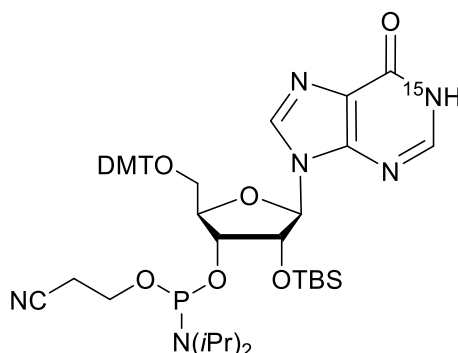
HRMS (ESI⁺): m/z berechnet für C₃₇H₄₅N₃¹⁵NO₇Si⁺ [M+H]⁺ 686.3022; gefunden: 686.3080.

HRMS (ESI⁻): m/z berechnet für C₃₇H₄₃N₃¹⁵NO₇Si⁻ [M-H]⁻ 684.2877; gefunden: 684.2943.

IR (ATR): ν (cm⁻¹) = 2925 (s), 2854 (m), 1692 (s), 1608 (w), 1586 (w), 1547 (w), 1509 (s), 1464 (m), 1416 (w), 1376 (w), 1343 (w), 1302 (w), 1251 (s), 1213 (w), 1176 (m), 1116 (w), 1092 (w), 1034 (m), 908 (w), 866 (w), 836 (s), 790 (m), 755 (w), 704 (w).

R_f (CH₂Cl₂:MeOH 15:1): 0.50.

3'-O-(Diisopropylamino-O-β-cyanoethoxyphosphino)-5'-O-(4,4'-dimethoxytrityl)-2'-(O-tert-butylidimethylsilyl)-[¹⁵N]-inosin (94**)**



Unter Ar-Atmosphäre wurde die DMT-geschützte Verbindung **93** (96 mg, 0.14 mmol, 1.0 eq.) in 5.0 mL trockenem CH₂Cl₂ gelöst und mit DIPEA (98 μ L, 0.56 mmol, 4.0 eq.) versetzt. Anschließend wurde das Reaktionsgemisch gründlich entgast (*freeze, pump, thaw* 3x) und schließlich tropfenweise mit CED-Cl (125 μ L, 0.56 mmol, 4.0 eq.) versetzt. Das Gemisch wurde anschließend 16 h bei Raumtemperatur verrührt und daraufhin direkt mittels Säulenchromatographie an Silicagel (CH₂Cl₂ \rightarrow CH₂Cl₂:MeOH 100:1 \rightarrow CH₂Cl₂:MeOH 80:1 \rightarrow CH₂Cl₂:MeOH 20:1, jeweils unter Zusatz von 1% NEt₃) aufgereinigt, um nach Entfernen der Lösungsmittel *in vacuo* und Gefriertrocknung des Rückstandes aus wasserfreiem Benzol (10 mL) 185 mg des Reaktionsproduktes **94** (126 μ mol, 90%, zwei Diastereomere am P) als NEt₃-Addukt (1.7 eq.) in Form eines farblosen Feststoffs zu erhalten.

¹H-NMR (400 MHz, CD₂Cl₂): δ/ppm = 8.02 – 7.91 (m, 2H, 2-H, 8-H), 7.51 – 7.18 (m, 9H, C-H_{Ar}(DMT)), 6.87 – 6.79 (m, 4H, C-H_{Ar}(DMT)), 6.00 – 5.93 (m, 1H, 1'-H), 4.92 – 4.87 (m, 1H, 2'-H), 4.45 – 4.32 (m, 2H, 3'-H, 4'-H), 4.21 – 3.81 (m, 2H, OCH₂CH₂CN), 3.78 (m, 6H, OCH₃), 3.70 – 3.42 (m, 3H, N-CH-(CH₃)₂, 5'-H), 3.32 (m, 1H, 5'-H), 2.78 (q, *J* = 7.2 Hz, 10H, N(CH₂CH₃)₃ aus NEt₃), 2.68 – 2.30 (m, 2H OCH₂CH₂CN), 1.28 – 1.02 (m, 30H, N-CH-(CH₃)₂ und N(CH₂CH₃)₃ aus NEt₃), 0.81 – 0.76 (m, 9H, SiC(CH₃)₃), 0.00 – -0.04 (m, 3H SiCH₃), -0.13 – -0.17 (m, 3H, SiCH₃).

³¹P-NMR (162 MHz, CD₂Cl₂): δ/ppm = 150.82, 149.01.

R_f (CH₂Cl₂:MeOH 20:1): 0.35.

Oligonukleotidsynthese

Oligonukleotid 1: 5'- ACU GGA CAI IUI CUC CGA GG

Oligonukleotid 2: 5'- CCU CGG AGU IUU UGU CCA GU

Die RNA-Oligonukleotidsynthese wurde unter Standardbedingungen mit einer Kartuschenkapazität von 1 µmol unter Verwendung der Phosphoramidite Bz-A, Ac-C, dmf-G und U (*LinkTechnologies*) durchgeführt. Die Phosphoramidite der kanonischen RNA-Nukleotide wurden als 0.1 M Lösung in MeCN, das Inosin-Phosphoramidit als 0.1 M Lösung in CH₂Cl₂ eingesetzt. Zur Bestimmung der eingesetzten Menge an Inosin-Phosphoramidit **94** wurde nur von 80% Ausbeute bezogen auf das eingesetzte Startmaterial ausgegangen.^[95] Zur Kopplung des Inosin-Phosphoramidits **94** wurde eine Kopplungszeit von 20 min jeweils unter Einsatz der doppelten Menge an Phosphoramidit verglichen zu den Phosphoramiditen der kanonischen RNA-Nukleoside verwendet. Die terminale DMT-Gruppe wurde nach der Oligonukleotidsynthese am Synthesizer abgespalten. Die Entschützung der Oligonukleotide erfolgte gemäß der Literaturvorschrift von *Dallmann et al.*,^[95] die abschließende Aufreinigung mittels *reversed-phase* HPLC bei einem Fluss von 5.0 mL/min und einem Gradienten von 0% auf 30% Puffer B in 45 min, wobei ein Puffersystem aus 0.1 M wässriger HNEt₃OAc-Lösung bei pH = 7 (Puffer A) und eine 0.1 M Lösung von HNEt₃OAc in 80% wässrigem MeCN bei pH = 7 (Puffer B) verwendet wurden. Die aufgereinigten Oligonukleotide wurden nach Entsalzen unter Verwendung von 0.025 µm VSWP Filtern (*Millipore*) gegen ddH₂O mit HPA-Matrix (3-Hydroxypicolinsäure) kokristallisiert und mittels MALDI-TOF analysiert.

7. Material und Methoden

Anmerkungen zu chemischen Synthesen und eingesetzten Chemikalien

Alle Reaktionen wurden, wenn nicht in der Synthesevorschrift anders angegeben, in unter Hochvakuum ausgeheizten Glasgeräten unter Anwendung der *Schlenk*-Technik unter Stickstoff-Atmosphäre durchgeführt. Die verwendeten Chemikalien wurden von den Firmen *Sigma-Aldrich*, *Eurisotop* (*Camebridge Isotope Laboratories*), *Fluka*, *TCI*, *abcr* oder *Carbosynth* in den entsprechenden Reinheitsgraden zur synthetischen Anwendung bezogen und ohne weitere Aufreinigung verwendet. Über Molekularsieb getrocknete und unter Schutzgasatmosphäre gelagerte Lösungsmittel wurden von der Firma *Acros Organics* bezogen und ausschließlich unter Schutzgas gehandhabt. Für Extraktionen, dünnschichtchromatographische Reaktionskontrollen und säulenchromatographische Aufreinigungen wurden Lösungsmittel technischen Reinheitsgrades verwendet, die vor ihrer Verwendung *in vacuo* destilliert wurden. Das Entfernen von Lösungsmitteln *in vacuo* wurde an einem Rotationsverdampfer *Laborota 4000* der Firma *Heidolph* durchgeführt, wobei die Wasserbadtemperatur falls in den Vorschriften nicht anders angegeben auf 40 °C eingestellt war.

Oligonukleotidsynthese

DNA- und RNA-Oligonukleotidsynthesen wurden an *ABI 394* DNA/RNA-Synthesizern (*Applied Biosystems*) unter Standardbedingungen durchgeführt. Die CPG-Kartuschen (*controlled pore glas*) sowie die Phosphoramidite der unmodifizierten Nukleoside wurden von *LinkTechnologies* bezogen. Für selbst-synthetisierte Phosphoramidite wurden gegebenenfalls Bedingungen der Oligonukleotid-Festphasensynthese angepasst. Diese Anpassungen sind im jeweiligen Kapitel zu finden.

Reaktionen unter Kohlenstoffmonoxid-Überdruck

Reaktionen unter Kohlenstoffmonoxid-Überdruck wurden in einem Druckautoklav bei einem Druck von 3.5 bar Kohlenstoffmonoxid durchgeführt. Für die Zugabe von Flüssigkeiten in das Reaktionsgemisch wurde eine Spritzenpumpe der Firma *KD Scientific* eingesetzt. Der Autoklav wurde vor der Reaktion im Ofen ausgeheizt, unter Argon-Strom abkühlen gelassen und schließlich mit Teilen des Reaktionsgemisches

vor dem Start der Reaktion dreimal mit 3.5 bar Kohlenstoffmonoxid für 15 min gespült.

Dünnschichtchromatographie (DC)

DC-Platten der Firma *Merck* (*F254*, Kieselgel auf Aluminium) wurden für Reaktionskontrollen und zur Überprüfung des Trennverhaltens von säulenchromatographischen Aufreinigungen verwendet. Substanzen wurden entweder durch ihre UV- oder Fluoreszenzaktivität unter einer UV-Lampe (254 nm und/oder 366 nm) und/oder durch Anfärben mit einer schwefelsauren Lösung von Molybdätophosphorsäure und Cer(II)-sulfat (2.5 g $\text{H}_3\text{PMo}_{12}\text{O}_{40}$ und 1.0 g CeSO_4 in 6 mL H_2SO_4 und 94 mL dd H_2O) und anschließendem Erhitzen auf 240 °C nachgewiesen.

Säulenchromatographie

Säulenchromatographische Aufreinigung erhaltener Rohprodukte wurde unter Stickstoffüberdruck gemäß der *Flash*-Methode durchgeführt, wobei als stationäre Phase *Kieselgel 60* der Firma *Merck* (Korngröße 0.063 – 0.200 nm) verwendet wurde. Passende Eluenten wurden über Dünnschichtchromatographie der Rohprodukte ermittelt.

High-Performance Liquid-Chromatography (HPLC)

Semi-präparative Aufreinigungen synthetisierter Verbindungen wurden an einem *Breeze*-System der Firma *Waters*, bestehend aus einer *1525 Binary HPLC Pump* und einem *2487 UV/Vis Detector*, oder mit einem *Agilent 1260 Infinity II* 400 bar-Pumpensystem mit einem *Agilent 1260 Infinity II VWD* bei 260 nm durchgeführt. Als stationäre Phase wurden *VP 250/10 Nucleosil 100-7 C18* oder *Nucleodur* (250/10 mm, C18ec, 5 µm) *reversed-phase* Chromatographiesäulen der Firma *Macherey & Nagel* bei einer Flussrate von 5.0 mL/min verwendet. Analytische HPLC-Untersuchungen wurden an einem *Waters* Separationsmodul *e2695* in Kombination mit einem *2487 UV/Vis Detector* (*Waters*) oder einem *Agilent 1260 Infinity II* 800 bar-*Flexible Prime*-Pumpensystem mit einem *Agilent 1260 Infinity II MWD* bei 260 nm (sofern nicht anders angegeben) durchgeführt. Als stationäre Phase wurden *120/3 Nucleosil C18*-Chromatographiesäulen der Firma *Macherey & Nagel* bei einer Flussrate von 0.5 mL/min oder 1.0 mL/min verwendet. Acetonitril wurde in HPLC-

Güte verwendet und bidestilliertes Wasser (ddH₂O) wurde mit Hilfe einer *arium® pro DI*-Anlage der Firma *sartorius stedim biotech* erzeugt. Falls gepufferte Eluentensysteme verwendet wurden, sind diese bei den jeweiligen Vorschriften genau charakterisiert. Für die Einstellung der pH-Werte von gepufferten Laufmittelsystemen wurde ein *MP 220* pH-Meter der Firma *Mettler Toledo* eingesetzt.

Gefriertrocknung

Zur Entfernung von Lösungsmitteln über Gefriertrocknung wurden *Alpha 2-4 LD plus*-Lyophilisen der Firma *Christ* verwendet. Lösungsmittel wurden vor der Gefriertrocknung in flüssigem Stickstoff eingefroren.

Kernresonanzspektroskopie (NMR-Spektroskopie)

Ein- und zweidimensionale NMR-Spektroskopie (¹H, ¹³C, ¹⁵N, ³¹P, COSY, HSQC, HMBC) wurde entweder selbstständig an einem *Bruker Ascend 400* NMR-Spektrometer durchgeführt oder von der Abteilung für NMR-Spektroskopie des Departments Chemie & Pharmazie der LMU München, namentlich von *Dr. D. Stephenson* und *C. Ober* durchgeführt. Bei diesen Messungen wurden als Spektrometer die Geräte *Bruker Avance III HD 400*, *Varian NMR-System 600* und *Bruker Avance III HD mit Cryo-Kopf 800* verwendet. Die chemischen Verschiebungen sind in ppm (*parts per million*) angegeben und werden relativ zum Signal des Lösungsmittels als jeweiligen internen Standard bezogen. Die Aufspaltungen der Signale gemäß der allgemeinen Nomenklatur mit s (Singulett), d (Dublett), t (Triplet), q (Quartett), p (Pentett), m (Multiplet) oder Kombinationen daraus angegeben. Kopplungskonstanten der Aufspaltungsmuster sind in *Hertz* angegeben.

Massenspektrometrie

Niederaufgelöste ESI-Massenspektrometrie zur Reaktionskontrolle wurde entweder per Direkteinspritzung oder über vorgeschaltete Flüssigchromatographie (LC-MS) selbstständig an einem *MSQ Plus* Massenspektrometer der Firma *Thermo Scientific* in Kombination mit einem *Dionex Ultimate3000* Flüssigchromatographie- und Einspritzsystem durchgeführt. Hochaufgelöste Massenspektrometrie wurde von der Abteilung für Massenspektrometrie des Departments für Chemie & Pharmazie der LMU München, namentlich von *Dr. W. Spahl* für die hochaufgelösten ESI-Massenspektren und *S. Kosak* für die hochaufgelösten EI-Massenspektren,

durchgeführt. Dabei wurden für die hochaufgelösten ESI-Messungen ein *LTQ-FT-ICR* der Firma *Finnigan* verwendet. MALDI-TOF-Massenspektrometrie wurde selbstständig am einem *Bruker Autoflex II* nach Entsalzen der Proben unter Verwendung von 0.025 μm VSWP Filtern (*Millipore*) gegen ddH₂O und anschließender Cokristallisation der Proben mit HPA-Matrix (3-Hydroxypicolinsäure) durchgeführt.

Infrarotspektroskopie

Infrarotspektroskopie wurde selbstständig an einem *Spectrum BX* FT-IR-Spektrometer der Firma *Perkin-Elmer* durchgeführt, wobei ein *DuraSampII R II* Messkopf der Firma *Smiths* und eine ATR-Einheit (*attenuated total reflection*-Einheit) aus Diamant verwendet wurden. Infrarotspektren der untersuchten Substanzen wurden entweder als Film der Substanz nach Verdampfen des Lösungsmittels einer Lösung der Probe oder durch Anpressen der festen Substanz auf die ATR-Einheit gemessen. Wellenzahlen der Signale sind in cm^{-1} angegeben und die Intensitäten der Signale durch die Bezeichnungen w (*weak*), m (*medium*), s (*strong*), charakterisiert.

UV/Vis-Spektroskopie

Für UV/Vis-spektroskopische Messungen wurde ein *JASCO V650* Photospektrometer, als Blindprobe zur Basislinienbestimmung vor den Messungen wurde das jeweilige Lösungsmittel der Probe verwendet.

8. Abkürzungsverzeichnis

A	Adenosin
ADAR	<i>Adenosine Deaminase acting on RNA</i>
a-fdC	6-Aza-5-formyl-2'-desoxycytidin
a-fdU	6-Aza-5-formyl-2'-desoxyuridin
AIBN	Azo(bisisobutyronitril)
AID	<i>Activation induced cytidine deaminase</i>
Am	2'-Methyladenosin
AML	akute myeloische Leukämie
APOBEC	<i>Apolipoprotein B m-RNA editing enzyme catalytic polypeptide</i>
Asn	Asparagin
Asp	Asparaginsäure
ATR	attenuated total reflection
Aza-C	5-Azacytidin
Aza-dC	5-Aza-2'-desoxycytidin
BER	Basenexzisionsreparatur
BSA	Bis(trimethylsilyl)-acetamid
C	Cytidin
cadC	5-Carboxy-2'-desoxycytidin
CAN	Cer(IV)-ammoniumnitrat
c-Aza-dC	carbazyklisches 5-Aza-2'-desoxycytidin
CDI	Carbonyldiimidazol
CEDCI	Cyanoethyl-diisopropyl-chloro-phosphoramidit
Cm	2'-Methylcytidin
COSY	Correlated Spectroscopy
CpG	dC-dG Dinkleotid
CuAAC	Kupfer katalysierte Alkin-Azid-Cycloaddition

dA	2'-Desoxyadenosin
dC	2'-Desoxycytidin
ddH ₂ O	bidestilliertes Wasser
DEMA	Diethoxymethylacetat
DIPEA	Diisopropylethylamin
(DIPSCI) ₂ O	Dichlor-1,1,3,3-tetraisopropyldisiloxan
dG	2'-Desoxyguanosin
DMAP	4-(Dimethylamino)-pyridin
DMF	Dimethylformamid
DMSO	Dimethylsulfoxid
DMT	4,4'-Dimethoxytrityl
DMTCI	4,4'-Dimethoxytritylchlorid
DMTOTf	4,4'-Dimethoxytrityltriflat
DNA	<i>desoxyribonucleic acid</i>
Dnmt	<i>DNA nucleotide methyltransferases</i>
dT	2'-Desoxythymidin
EDC-HCl	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimid Hydrochlorid
EDTA	<i>Ethylenediaminetetraacetic acid</i>
ESI	Elektronenspray-Ionisation
fdC	5-Formyl-2'-desoxycytidin
F-dC	2'-(<i>R</i>)-Fluor-2'-desoxycytidin
fdU	5-Formyl-2'-desoxyuridin
F-fdC	5-Formyl-2'-(<i>R</i>)-fluor-2'-desoxycytidin
F-mdC	5-Methyl-2'-(<i>R</i>)-fluor-2'-desoxycytidin
G	Guanosin
GalQ	β-Galaktosylqueuosin
Gm	2'-Methylguanosin

HATU	O-(7-Azabenzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphat
HBTU	2-(1H-Benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphat
HDAC	Histon-Deacetylase
HEPES	4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid
His	Histidin
HMBC	<i>heteronuclear multiple bond correlation</i>
hmdC	5-Hydroxymethyl-2'-desoxycytidin
HMDS	Hexamethyldisilazan
HOBt	1-hydroxybenzotriazol
HPA	3-Hydroxypicolinic acid
HPLC	<i>high-performance liquid chromatography</i>
HRMS	<i>high resolution mass spectrometry</i>
hSirt1	humanes Sirtuin1
HSQC	<i>Heteronuclear Single Quantum Coherence</i>
I	Inosin
iNGNs	<i>small molecule-inducible neurogenin iPS-line</i>
IPD	<i>interpulse duration</i>
iPSCs	induzierte pluripotente Stammzellen
IR	Infrarotspektroskopie
KO	<i>knockout</i>
LC-MS	<i>liquid-chromatography coupled mass spectrometry</i>
LMU	Ludwigs-Maximilians Universität
lncRNA	<i>long non-coding RNA</i>
m5C	5-Methylcytidin
MALDI	<i>matrix assisted laser desorption ionisation</i>
ManQ	Mannosylqueuosin

MBD4	<i>methyl CpG binding domain 4</i>
mCPBA	<i>meta</i> -Chlorperbenzoesäure
mdC	5-Methyl-2'-desoxycytidin
MDS	myelodysplastisches Syndrom
MeCN	Acetonitril
mESCs	murine embryonale Stammzellen
miRNA	<i>micro RNA</i>
mRNA	<i>messenger RNA</i>
MS	Massenspektrometrie
NAM	Nikotinamid
NHS	<i>N</i> -Hydroxysuccinimid
NMI	<i>N</i> -Methylimidazol
NMR	<i>nuclear magnetic resonance</i>
NOESY	<i>nuclear Overhauser enhancement spectroscopy</i>
PacBio	<i>Pacific Biosciences</i>
PCR	<i>polymerase chain reaction</i>
PDO	Propan-1,3-diol
PEG	Polyethylenglykol
ppm	<i>parts per million</i>
Py	Pyridin
Q	Queuosin
R _f	Retentionsfaktor
RNA	<i>ribonucleic acid</i>
rRNA	<i>ribosomal RNA</i>
RT	Raumtemperatur
SAM	S-Adenosylmethionin
siRNA	<i>small interfering RNA</i>

SMRT	<i>single molecule real time sequencing</i>
SMUG1	<i>single strand selective monofunctional uracil-DNA glycosylase</i>
S _N 2	nukleophile Substitution 2. Ordnung
TBAI	Tetrabutylammoniumiodid
TBSCI	<i>tert</i> -Butyldimethylsilylchlorid
TDG	Thymin-DNA Glykosylase
TET	<i>ten-eleven translocation</i>
TFA	<i>trifluoroacetic acid</i> , Trifluoressigsäure
THF	Tetrahydrofuran
TMSCI	Chlortrimethylsilan
TMSOMe	Methoxytrimethylsilan
TMSOTf	Trimethylsilyltrifluormethansulfonat
TOF	<i>time of flight</i>
TPSCI	2,4,6-Triisopropylbezolsulfonylchlorid
TQ	<i>triple quadrupole</i>
tRNA	<i>transfer RNA</i>
TU München	Technische Universität München
Tyr	Tyrosin
U	Uridin
UHPLC	<i>ultra high-performance liquid chromatography</i>
UV/Vis	Ultraviolette und sichtbares Licht

9. Literaturverzeichnis

- [1] R. Holliday, *Dev. Genet.* **1994**, 15, 453457.
- [2] A. Breiling, F. Lyko, *Epigenetics and Chromatin* **2015**, 8, 1–9.
- [3] R. J. Klose, A. P. Bird, *Trends Biochem. Sci.* **2006**, 31, 89–97.
- [4] B. D. Strahl, C. D. Allis, *Nature* **2000**, 403, 41–45.
- [5] T. Pfaffeneder, B. Hackner, M. Truß, M. Münzel, M. Müller, C. A. Deiml, C. Hagemeyer, T. Carell, *Angew. Chemie* **2011**, 123, 7146–7150.
- [6] J. A. Law, S. E. Jacobsen, *Nat. Rev. Genet.* **2010**, 11, 204–220.
- [7] J. Doskočil, F. Šorm, *Biochim. Biophys. Acta - Spec. Sect. Nucleic Acids Relat. Subj.* **1962**, 55, 953–959.
- [8] P. A. Jones, *Nat. Rev. Genet.* **2012**, 13, 484–492.
- [9] S. C. Wu, Y. Zhang, *Nat. Rev. Mol. Cell Biol.* **2010**, 11, 607–620.
- [10] T. Baubec, D. Schübeler, *Curr. Opin. Genet. Dev.* **2014**, 25, 85–92.
- [11] Z. D. Smith, A. Meissner, *Nat. Rev. Genet.* **2013**, 14, 204–220.
- [12] B. Panning, R. Jaenisch, *Genes Dev.* **1996**, 10, 1991–2002.
- [13] J. L. Swain, T. A. Stewart, P. Leder, *Cell* **1987**, 50, 719–727.
- [14] E. Li, C. Beard, R. Jaenisch, *Nature* **1993**, 366, 362–365.
- [15] A. Jeltsch, *Curr. Top. Biochem.* **2002**, 274–293.
- [16] C. P. Walsh, J. R. Chaillet, T. H. Bestor, *Nat. Genet.* **1998**, 20, 116–117.
- [17] A. P. Feinberg, B. Tycko, *Nat. Rev. Cancer* **2004**, 4, 143–153.
- [18] C. B. Santos-Rebouças, M. M. G. Pimentel, *Eur. J. Hum. Genet.* **2007**, 15, 10–17.
- [19] I. P. Pogribny, F. A. Beland, *Cell. Mol. Life Sci.* **2009**, 66, 2249–2261.
- [20] R. Goyal, R. Reinhardt, A. Jeltsch, *Nucleic Acids Res.* **2006**, 34, 1182–1188.
- [21] M. G. Goll, T. H. Bestor, *Annu. Rev. Biochem.* **2005**, 74, 481–514.
- [22] R. Z. Jurkowska, T. P. Jurkowski, A. Jeltsch, *ChemBioChem* **2011**, 12, 206–222.
- [23] R. Jüttermann, E. Li, R. Jaenisch, *Proc. Natl. Acad. Sci.* **1994**, 91, 11797 LP – 11801.
- [24] P. A. Jones, *Pharmacol. Ther.* **1985**, 28, 17–27.
- [25] S. Gabbara, A. S. Bhagwat, *Biochem. J.* **1995**, 307 (Pt 1, 87–92.
- [26] A. Čihák, *Oncology* **1974**, 30, 405–422.
- [27] J. K. Christman, N. Mendelsohn, D. Herzog, N. Schneiderman, *Cancer Res.* **1983**, 43, 763 LP – 769.
- [28] M. Daskalakis, T. T. Nguyen, C. Nguyen, P. Guldberg, G. Köhler, P. Wijermans, P. A. Jones, M. Lübbert, *Blood* **2002**, 100, 2957–2964.
- [29] L. R. Silverman, J. F. Holland, R. S. Weinberg, B. P. Alter, R. B. Davis, R. R. Ellison, E. P. Demakos, C. J. Cornell, R. W. Carey, C. Schiffer, *Leukemia* **1993**, 7 Suppl 1, 21–29.
- [30] H. Kantarjian, J.-P. J. Issa, C. S. Rosenfeld, J. M. Bennett, M. Albitar, J. DiPersio, V. Klimek, J. Slack, C. de Castro, F. Ravandi, et al., *Cancer* **2006**, 106, 1794–1803.
- [31] C. Stresemann, F. Lyko, *Int. J. Cancer* **2008**, 123, 8–13.
- [32] T. M. Wildenhof, S. Schiffers, F. R. Traube, P. Mayer, T. Carell, *Angew. Chemie* **2019**, 131, 13118–13121.
- [33] D. K. Rogstad, J. L. Herring, J. A. Theruvathu, A. Burdzy, C. C. Perry, J. W. Neidigh, L. C. Sowers, *Chem. Res. Toxicol.* **2009**, 22, 1194–1204.
- [34] S. E. Walker, L. F. Charbonneau, S. Law, C. Earle, *Can. J. Hosp. Pharm.* **2012**, 65, DOI 10.4212/cjhp.v65i5.1172.
- [35] Y. Kato, S. Ozawa, C. Miyamoto, Y. Maehata, A. Suzuki, T. Maeda, Y. Baba,

- Cancer Cell Int.* **2013**, *13*, 89.
- [36] M. Tahiliani, K. P. Koh, Y. Shen, W. A. Pastor, H. Bandukwala, Y. Brudno, S. Agarwal, L. M. Iyer, D. R. Liu, L. Aravind, et al., *Science* **2009**, *324*, 930–935.
 - [37] S. Ito, L. Shen, Q. Dai, S. C. Wu, L. B. Collins, J. A. Swenberg, C. He, Y. Zhang, *Science* **2011**, *333*, 1300–1303.
 - [38] S. Ito, A. C. Dalessio, O. V. Taranova, K. Hong, L. C. Sowers, Y. Zhang, *Nature* **2010**, *466*, 1129–1133.
 - [39] C. Loenarz, C. J. Schofield, *Chem. Biol.* **2009**, *16*, 580–583.
 - [40] S. Kriaucionis, N. Heintz, *Science* **2009**, *324*, 929–930.
 - [41] Y.-F. He, B.-Z. Li, Z. Li, P. Liu, Y. Wang, Q. Tang, J. Ding, Y. Jia, Z. Chen, L. Li, et al., *Science* **2011**, *333*, 1303 – 1307.
 - [42] H. Wu, A. C. D'Alessio, S. Ito, K. Xia, Z. Wang, K. Cui, K. Zhao, Y. Eve Sun, Y. Zhang, *Genes Dev.* **2011**, *25*, 679–684.
 - [43] M. Münzel, D. Globisch, T. Carell, *Angew. Chemie - Int. Ed.* **2011**, *50*, 6460–6468.
 - [44] D. Globisch, M. Münzel, M. Müller, S. Michalakis, M. Wagner, S. Koch, T. Brückl, M. Biel, T. Carell, *PLoS One* **2010**, *5*, e15367.
 - [45] M. Münzel, D. Globisch, T. Brückl, M. Wagner, V. Welzmler, S. Michalakis, M. Müller, M. Biel, T. Carell, *Angew. Chem. Int. Ed.* **2010**, *49*, 5375–5377.
 - [46] M. Wagner, J. Steinbacher, T. F. J. Kraus, S. Michalakis, B. Hackner, T. Pfaffeneder, A. Perera, M. Müller, A. Giese, H. A. Kretzschmar, et al., *Angew. Chem. Int. Ed.* **2015**, *54*, 12511–12514.
 - [47] S. Kriaucionis, N. Heintz, *Science* **2009**, *324*, 929 – 930.
 - [48] T. Carell, M. Q. Kurz, M. Müller, M. Rossa, F. Spada, *Angew. Chem. Int. Ed.* **2018**, *57*, 4296–4312.
 - [49] M. Bachman, S. Uribe-Lewis, X. Yang, H. E. Burgess, M. Iurlaro, W. Reik, A. Murrell, S. Balasubramanian, *Nat. Chem. Biol.* **2015**, *11*, 555–557.
 - [50] A. Schön, E. Kaminska, F. Schelter, E. Ponkkonen, E. Korytiaková, S. Schiffers, T. Carell, *Angew. Chem. Int. Ed.* **2020**, *59*, 5591–5594.
 - [51] S. Schiesser, B. Hackner, T. Pfaffeneder, M. Müller, C. Hagemeier, M. Truss, T. Carell, *Angew. Chem. Int. Ed.* **2012**, *51*, 6516–6520.
 - [52] S. Cortellino, J. Xu, M. Sannai, R. Moore, E. Caretti, A. Cigliano, M. Le Coz, K. Devarajan, A. Wessels, D. Soprano, et al., *Cell* **2011**, *146*, 67–79.
 - [53] R. Rahimoff, O. Kosmatchev, A. Kirchner, T. Pfaffeneder, F. Spada, V. Brantl, M. Müller, T. Carell, *J. Am. Chem. Soc.* **2017**, *139*, 10359–10364.
 - [54] P. Fortini, E. Dogliotti, *DNA Repair (Amst)*. **2007**, *6*, 398–409.
 - [55] R. Métivier, R. Gallais, C. Tiffocche, C. Le Péron, R. Z. Jurkowska, R. P. Carmouche, D. Ibberson, P. Barath, F. Demay, G. Reid, et al., *Nature* **2008**, *452*, 45–50.
 - [56] K. A. Haushalter, P. T. Stukenberg, M. W. Kirschner, G. L. Verdine, *Curr. Biol.* **1999**, *9*, 174–185.
 - [57] R. J. Boorstein, A. Cummings, D. R. Marenstein, M. K. Chan, Y. Ma, T. A. Neubert, S. M. Brown, G. W. Teebor, *J. Biol. Chem.* **2001**, *276*, 41991–41997.
 - [58] T. Pfaffeneder, F. Spada, M. Wagner, C. Brandmayr, S. K. Laube, D. Eisen, M. Truss, J. Steinbacher, B. Hackner, O. Kotljaraova, et al., *Nat. Chem. Biol.* **2014**, *10*, 574–581.
 - [59] K. Iwan, R. Rahimoff, A. Kirchner, F. Spada, A. S. Schröder, O. Kosmatchev, S. Ferizaj, J. Steinbacher, E. Parsa, M. Müller, et al., *Nat. Chem. Biol.* **2018**, *14*, 72–78.
 - [60] M. Münzel, D. Globisch, T. Carell, *Angew. Chem. Int. Ed.* **2011**, *50*, 6460–6468.

- [61] A. S. Schröder, O. Kotljarova, E. Parsa, K. Iwan, N. Raddaoui, T. Carell, *Org. Lett.* **2016**, *18*, 4368–4371.
- [62] E.-A. Raiber, R. Hardisty, P. van Delft, S. Balasubramanian, *Nat. Rev. Chem.* **2017**, *1*, DOI 10.1038/s41570-017-0069.
- [63] C. X. Song, C. Yi, C. He, *Nat. Biotechnol.* **2012**, *30*, 1107–1116.
- [64] J. Korlach, S. W. Turner, *Curr. Opin. Struct. Biol.* **2012**, *22*, 251–261.
- [65] M. Ronaghi, *Genome Res.* **2001**, *11*, 3–11.
- [66] R. Shapiro, R. E. Servis, M. Welcher, *J. Am. Chem. Soc.* **1970**, *92*, 422–424.
- [67] H. Hayatsu, Y. Wataya, K. Kai, S. Iida, *Biochemistry* **1970**, *9*, 2858–2865.
- [68] E. Kriukiene, Z. Liutkevičiute, S. Klimašauskas, *Chem. Soc. Rev.* **2012**, *41*, 6916–6930.
- [69] M. J. Booth, G. Marsico, M. Bachman, D. Beraldi, S. Balasubramanian, *Nat. Chem.* **2014**, *6*, 435–440.
- [70] C. X. Song, K. E. Szulwach, Q. Dai, Y. Fu, S. Q. Mao, L. Lin, C. Street, Y. Li, M. Poidevin, H. Wu, et al., *Cell* **2013**, *153*, 678–691.
- [71] M. J. Booth, M. R. Branco, G. Ficz, D. Oxley, F. Krueger, *Science* **2012**, *336*, 934–937.
- [72] X. Lu, C. Song, K. Szulwach, Z. Wang, P. Weidenbacher, P. Jin, C. He, *J. Am. Chem. Soc.* **2013**, *135*, 9315–9317.
- [73] B. Xia, D. Han, X. Lu, Z. Sun, A. Zhou, Q. Yin, H. Zeng, M. Liu, X. Jiang, W. Xie, et al., *Nat. Methods* **2015**, *12*, 1047–1050.
- [74] X. Wu, A. Inoue, T. Suzuki, Y. Zhang, *Genes Dev.* **2017**, *31*, 511–523.
- [75] A. Rhoads, K. F. Au, *Genom. Proteom. Bioinf.* **2015**, *13*, 278–289.
- [76] T. A. Clark, X. Lu, K. Luong, Q. Dai, M. Boitano, S. W. Turner, C. He, J. Korlach, *BMC Biol.* **2013**, *11*, DOI 10.1186/1741-7007-11-4.
- [77] L. Chavez, Y. Huang, K. Luong, S. Agarwald, L. M. Iyer, W. A. Pastor, V. K. Hench, S. A. Frazier-Bowers, E. Korol, S. Liuh, et al., *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, E5149–E5158.
- [78] C. X. Song, T. A. Clark, X. Y. Lu, A. Kislyuk, Q. Dai, S. W. Turner, C. He, J. Korlach, *Nat. Methods* **2012**, *9*, 75–77.
- [79] P. Boccaletto, M. A. Machnicka, E. Purta, P. Piątkowski, B. Bagiński, T. K. Wirecki, V. de Crécy-Lagard, R. Ross, P. A. Limbach, A. Kotter, et al., *Nucleic Acids Res.* **2017**, *46*, D303–D307.
- [80] W. A. Cantara, P. F. Crain, J. Rozenski, J. A. McCloskey, K. A. Harris, X. Zhang, F. A. P. Vendeix, D. Fabris, P. F. Agris, *Nucleic Acids Res.* **2010**, *39*, D195–D201.
- [81] L. He, G. J. Hannon, *Nat. Rev. Genet.* **2004**, *5*, 522–531.
- [82] M. Guttman, J. L. Rinn, *Nature* **2012**, *482*, 339–346.
- [83] T. R. Cech, *Gene* **1993**, *135*, 33–36.
- [84] P. A. Limbach, P. F. Crain, J. A. McCloskey, *Nucleic Acids Res.* **1994**, *22*, 2183–2196.
- [85] F. Klepper, E.-M. Jahn, V. Hickmann, T. Carell, *Angew. Chem. Int. Ed.* **2007**, *46*, 2325–2327.
- [86] C. Brandmayr, M. Wagner, T. Brückl, D. Globisch, D. Pearson, A. C. Kneutinger, V. Reiter, A. Hienzs, S. Koch, I. Thoma, et al., *Angew. Chem. Int. Ed.* **2012**, *51*, 11162–11165.
- [87] D. Globisch, D. Pearson, A. Hienzs, T. Brückl, M. Wagner, I. Thoma, P. Thumbs, V. Reiter, A. C. Kneutinger, M. Müller, et al., *Angew. Chem. Int. Ed.* **2011**, *50*, 9739–9742.
- [88] T. Suzuki, T. Suzuki, *Nucleic Acids Res.* **2014**, *42*, 7346–7357.
- [89] A. Costa, J. P. de Barros, G. Keith, W. Baranowski, J. Desgrès, *J. Chromatogr.*

- B. Analyt. Technol. Biomed. Life Sci.* **2004**, 801, 237–247.
- [90] H. Grosjean, M. Sprinzl, S. Steinberg, *Biochimie* **1995**, 77, 139–141.
- [91] H. Kasai, K. Nakanishi, R. D. Macfarlane, D. F. Torgerson, Z. Ohashi, J. A. McCloskey, H. J. Gross, S. Nishimura, *J. Am. Chem. Soc.* **1976**, 98, 5044–5046.
- [92] F. Harada, S. Nishimura, *Biochemistry* **1972**, 11, 301–308.
- [93] N. Okada, N. Shindo-Okada, S. Nishimura, *Nucleic Acids Res.* **1977**, 4, 415–423.
- [94] N. Okada, S. Nishimura, *Nucleic Acids Res.* **1977**, 4, 2931–2937.
- [95] A. Dallmann, A. V. Beribisky, F. Gnerlich, M. Rübhelke, S. Schiesser, T. Carell, M. Sattler, *Chem. Eur. J.* **2016**, 22, 15350–15359.
- [96] B. L. Bass, H. Weintraub, *Cell* **1988**, 55, 1089–1098.
- [97] K. Nishikura, *Nat. Rev. Mol. Cell Biol.* **2016**, 17, 83–96.
- [98] I. Alseth, B. Dalhus, M. Bjørås, *Curr. Opin. Genet. Dev.* **2014**, 26, 116–123.
- [99] S. Farajollahi, S. Maas, *Trends Genet.* **2010**, 26, 221–230.
- [100] S. M. Rueter, T. R. Dawson, R. B. Emeson, *Nature* **1999**, 399, 75–80.
- [101] N. Paz-Yaacov, L. Bazak, I. Buchumenski, H. T. Porath, M. Danan-Gotthold, B. A. Knisbacher, E. Eisenberg, E. Y. Levanon, *Cell Rep.* **2015**, 13, 267–276.
- [102] L. Chen, Y. Li, C. H. Lin, T. H. M. Chan, R. K. K. Chow, Y. Song, M. Liu, Y.-F. Yuan, L. Fu, K. L. Kong, et al., *Nat. Med.* **2013**, 19, 209–216.
- [103] L. Han, L. Diao, S. Yu, X. Xu, J. Li, R. Zhang, Y. Yang, H. M. J. Werner, A. K. Eterovic, Y. Yuan, et al., *Cancer Cell* **2015**, 28, 515–528.
- [104] P. Thumbs, Dissertation, **2013**.
- [105] P. K. Chang, *J. Org. Chem.* **1961**, 26, 1118–1120.
- [106] M. Kögler, R. Busson, S. De Jonghe, J. Rozenski, K. Van Belle, T. Louat, H. Munier-Lehmann, P. Herdewijn, *Chem. Biodivers.* **2012**, 9, 536–556.
- [107] J. Gut, *Adv. Heterocycl. Chem.* **1963**, 1, 189–251.
- [108] M. Bobek, J. Farkaš, F. Šorm, *Collect. Czechoslov. Chem. Commun.* **1967**, 32, 3581–3586.
- [109] M. Bobek, J. Farkaš, F. Šorm, *Collect. Czechoslov. Chem. Commun.* **1965**, 30, 3134–3139.
- [110] S. Schiffers, Dissertation, **2019**.
- [111] H. Hashimoto, J. E. Pais, N. Dai, I. R. Corrêa Jr, X. Zhang, Y. Zheng, X. Cheng, *Nucleic Acids Res.* **2015**, 43, 10713–10721.
- [112] H. Hashimoto, J. E. Pais, X. Zhang, L. Saleh, Z.-Q. Fu, N. Dai, I. R. Corrêa, Y. Zheng, X. Cheng, *Nature* **2014**, 506, 391–395.
- [113] L. Hu, Z. Li, J. Cheng, Q. Rao, W. Gong, M. Liu, Y. G. Shi, J. Zhu, P. Wang, Y. Xu, *Cell* **2013**, 155, 1545–1555.
- [114] R. Rahimoff, Dissertation, **2018**.
- [115] A. Kirchner, Dissertation, **2020**.
- [116] K. Iwan, Dissertation, **2018**.
- [117] Y. Saito, T. A. Zevaco, L. A. Agrofoglio, *Tetrahedron* **2002**, 58, 9593–9603.
- [118] M.-Z. Zhao, Y.-W. Zhang, F. Yuan, Y. Deng, J.-X. Liu, Y.-L. Zhou, X.-X. Zhang, *Talanta* **2015**, 144, 992–997.
- [119] Y. Yu, F. Yuan, X. H. Zhang, M. Z. Zhao, Y. L. Zhou, X. X. Zhang, *Anal. Chem.* **2019**, 91, 13047–13053.
- [120] F. R. Traube, S. Schiffers, K. Iwan, S. Kellner, F. Spada, M. Müller, T. Carell, *Nat. Protoc.* **2019**, 14, 283–312.
- [121] J. Seidel, C. Klockenbusch, D. Schwarzer, *ChemBioChem* **2016**, 17, 398–402.
- [122] S. Schiesser, T. Pfaffeneder, K. Sadeghian, B. Hackner, B. Steigenberger, A. S. Schröder, J. Steinbacher, G. Kashiwazaki, G. Höfner, K. T. Wanner, et al., *J.*

- Am. Chem. Soc.* **2013**, *135*, 14593–14599.
- [123] H. Jing, H. Lin, *Chem. Rev.* **2015**, *115*, 2350–2375.
- [124] N. Kojima, T. Takebayashi, A. Mikami, E. Ohtsuka, Y. Komatsu, *J. Am. Chem. Soc.* **2009**, *131*, 13208–13209.
- [125] H. Langhals, A. Obermeier, *European J. Org. Chem.* **2008**, 6144–6151.
- [126] A. T. M. e Silva, A. L. C. Maia, J. de Oliveira Silva, A. L. B. de Barros, D. C. F. Soares, M. T. Q. de Magalhães, R. José Alves, G. A. Ramaldes, *Carbohydr. Res.* **2018**, *465*, 52–57.
- [127] C. Marculescu, H. Kossen, R. E. Morgan, P. Mayer, S. A. Fletcher, B. Tolner, K. A. Chester, L. H. Jones, J. R. Baker, *Chem. Commun.* **2014**, *50*, 7139–7142.
- [128] N. Vurgun, R. F. Gómez-Biagi, M. Nitz, *Chem. Eur. J.* **2016**, *22*, 1062–1069.
- [129] J.-P. Couturier, M. Sütterlin, A. Laschewsky, C. Hettrich, E. Wischerhoff, *Angew. Chemie Int. Ed.* **2015**, *54*, 6641–6644.
- [130] K. Hashida, R. Makino, S. Ohara, *Holzforschung* **2009**, *63*, DOI 10.1515/HF.2009.043.
- [131] A. Alam, S. Tsuboi, *Tetrahedron* **2007**, *63*, 10454–10465.
- [132] B. L. Kam, N. J. Oppenheimer, *J. Org. Chem.* **1981**, *46*, 3268–3272.
- [133] R. Y. THOMSON, G. T. SCOTTO, G. B. BROWN, *J. Biol. Chem.* **1962**, *237*, 3510–3512.
- [134] P. J. Thornton, H. Kadri, A. Miccoli, Y. Mehellou, *J. Med. Chem.* **2016**, *59*, 10400–10410.
- [135] M. J. Sofia, *Antivir. Chem. Chemother.* **2011**, *22*, 23–49.
- [136] M. J. Sofia, D. Bao, W. Chang, J. Du, D. Nagarathnam, S. Rachakonda, P. G. Reddy, B. S. Ross, P. Wang, H. R. Zhang, et al., *J. Med. Chem.* **2010**, *53*, 7202–7218.
- [137] U. Pradere, E. C. Garnier-Amblard, S. J. Coats, F. Amblard, R. F. Schinazi, *Chem. Rev.* **2014**, *114*, 9154–9218.
- [138] M. S. Davey, R. Malde, R. C. Mykura, A. T. Baker, T. E. Taher, C. S. Le Duff, B. E. Willcox, Y. Mehellou, *J. Med. Chem.* **2018**, *61*, 2111–2117.
- [139] M. Uchiyama, Y. Aso, R. Noyori, Y. Hayakawa, *J. Org. Chem.* **1993**, *58*, 373–379.
- [140] C. Gardelli, B. Attenni, M. Donghi, M. Meppen, B. Pacini, S. Harper, A. Di Marco, F. Fiore, C. Giuliano, V. Pucci, et al., *J. Med. Chem.* **2009**, *52*, 5394–5407.
- [141] M. L. Jain, Y. P. Tsao, N. L. Ho, J. W. Cheng, *J. Org. Chem.* **2001**, *66*, 6472–6475.
- [142] A. R. Pagano, W. M. Lajewski, R. A. Jones, *J. Am. Chem. Soc.* **1995**, *117*, 11669–11672.

10. Danksagung

An erster Stelle möchte ich meinem Doktorvater *Prof. Dr. Thomas Carell* für die interessante Themenstellung und ausgezeichnete Betreuung im Rahmen von vielen hilfreichen wissenschaftlichen Diskussionen sowie für das mir gleichzeitig entgegengebrachten Vertrauen in meine Arbeit danken.

Für die Übernahme des Zweitgutachtens danke ich ganz herzlich *Dr. Sabine Schneider*.

Für die Unterstützung im Hinblick auf die Erledigung aller anstehenden bürokratischen Aufgaben danke ich Frau *Slava Gärtner*.

Dr. Markus Müller möchte ich ganz herzlich für die vielen wissenschaftlichen Diskussionen und Ratschläge, die mir stets weitergeholfen haben, danken.

Kerstin Kurz, Kristof Hufnagel und *Luis de la Osa de da Rosa* sei für deren stetige und freundliche Unterstützung in vielerlei Hinsichten gedankt.

Selbstverständlich möchte ich für das überprüfende Lesen dieser Dissertation *Stefan Wiedemann, Florian Schelter* und *Dr. Markus Müller* danken.

Allen Mitgliedern des Arbeitskreises danke ich für die gute Arbeitsatmosphäre während meiner Zeit dort.

Auch den Studenten und Praktikanten, die mit mir im Rahmen von Laborpraktika, Bachelor-Arbeiten oder F-Praktika gearbeitet haben, möchte ich für deren Mithilfe danken.

Insbesondere *Markus Hillmeier*, den ich im Rahmen seiner Master-Arbeit betreuen durfte, möchte ich für seine harte Arbeit und seine beeindruckende Gabe zur Begeisterung für die organische Synthesechemie danken.

Auch den Mitgliedern des „*HPLC-Serviceteams*“, dessen Mitglied ich mehrere Jahre sein durfte, namentlich *Matthias Kurz, Samuele Stazzoni, Stefan Wiedemann* und *Florian Schelter*, möchte ich für die vielen, teils unterhaltsamen und abenteuerlichen Stunden, die wir im Rahmen dieser Aufgabe verbracht haben, danken.

Ganz besonders möchte ich der Schafkopfrunde um *Stefan Wiedemann* („Des wollt ich jetzt einfach net wegschmeißen...“), *Florian Schelter* („Der Erdbeerschorsch“),

Markus Hillmeier („A Grasl, des hat jeder...“) und dem besten nicht-bayerischen Schafkopfspieler *Felix Hagelskamp* („Was, wenn ich da jetzt frei bin...?“) für die vielen unterhaltsamen Abende danken.

Natürlich möchte ich auch meiner Familie für die stetige Unterstützung während meines Studiums und meiner Doktorarbeit danken. Zuletzt, und am meisten, möchte ich meiner Lebenspartnerin und Mutter unseres gemeinsamen, kommenden Kindes *Eveliina Ponkkonen* danken. Ohne Deine bedingungslose Unterstützung, allgegenwärtige Aufmunterung und Ermutigung, die Du mir immer gezeigt und entgegengebracht hast, wäre diese Arbeit nicht möglich gewesen. Auch Deine Stärke und Willenskraft haben mich stets inspiriert, Dein kritisches Hinterfragen meiner Arbeit hat mich immer vorangebracht. Für jeden Tag, den wir bisher zusammen verbringen durften und in Zukunft miteinander verbringen dürfen, bin ich unendlich dankbar. *Kiitos kulta!*